

**ZAMM**

Zeitschrift für Angewandte Mathematik und Mechanik

Founded by Richard von Mises in 1921

Edited in cooperation with Universität Potsdam  
and Gesellschaft für Angewandte Mathematik und Mechanik e. V. (GAMM)

Editors-in-Chief: E. Kreuzer, R. Mennicken, H. Schwetlick

**DISTRIBUTION STATEMENT A**  
Approved for Public Release  
Distribution Unlimited

**GAMM 2000**

Annual Meeting

Minisymposia

University of Göttingen, Germany  
April 2–7, 2000

Edited by Gerd E. A. Meier, DLR Göttingen

20010927 020

 **WILEY-VCH**

Volume 81 · 2001  
Supplement 1

## To our Authors

ZAMM publishes original papers (main articles and short communications), the lectures presented at the annual GAMM conferences, and survey articles (requested by the Editorial Office of ZAMM). Papers in English (preferably) to be submitted for publication should be sent to the address at the bottom of this page (master and one copy), in the case of original papers together with an explicit assurance that they have not been published previously in, or are submitted to, another public journal. Authors are held fully responsible for all information included in their paper and should have one copy available for correction purposes. They get two sets of page proofs, one of them has to be returned within one week. Authors may be charged for extensive corrections in the text; corrections in the figures are not possible.

Manuscripts may be submitted in typewritten form or on disk together with a printout that is identical with the text on the disk. Manuscripts should be one-sided, double-spaced with wide margins at both sides, each sheet numbered, beginning with authors' names and the title of the paper. Extra types (like bold-faced letters etc.) should be marked or printed. Please, distinguish 1, l, e; v, w,  $\omega$ ; o, O, 0 (zero); k, K,  $\kappa$ ; i;  $\iota$ ; x,  $\times$  (cross),  $\chi$  etc. **Formulas** must be typed clearly legible (never use pencil). Formulas used in the text have to be numbered at the right margin. Use abbreviations for long terms, "exp" instead of "e" for extensive exponents, bold-faced letters for vectors etc. instead of arrows; avoid hierarchies of indices and exponents if possible. The end of a proof should be marked.

When preparing your ZAMM article with LATEX, please use the special class file **zamm.cls** (for LATEX 2e) or style file **zamm.sty** (for LATEX 2.09), respectively. Please, download these files from the internet site <http://www.wiley-vch.de/berlin/journals/zamm> (or contact the editorial office).

**Figures** should be submitted on separate sheets, drawn with Indian ink and numbered consecutively. Figure captions, tables, and footnotes (if not avoidable) should be inserted in the final place in the text.

**References** should be numbered in brackets throughout the text and collected according to these numbers in the appended bibliography (References). Each reference should contain Author(s), title, and complete bibliographical coordinates following the subsequent examples.

Books (Reports, Preprints etc.):

IRONS, B. M.; SHRIVE, N. G.: Numerical methods in engineering and applied science – Numbers are fun. Ellis Horwood Series in Mathematics and Its Applications. Ellis Horwood Ltd., Chichester 1987.

Articles in periodicals:

KASSEM, S. E.: Wave source potentials for two superposed fluids, each of finite depth. *Indian J. Pure Appl. Math.* **18** (1987) 2, 186–193.

Articles in Proceedings etc.:

LABISCH, F. K.: Some remarks on the morphology of nonunique solutions in nonlinear elastostatics. In: KÜPPER, T.; SEYDEL, R.; TROGER, H. (eds.): Bifurcation: Analysis, algorithms, applications. Proc. Conf. at the Univ. of Dortmund, August 18–22, 1988. *Internat. Ser. of Numer. Mathematics*, Vol. 79, Birkhäuser Verlag, Basel–Boston–Stuttgart 1987, pp. 177–184.

To each original paper and review article, a **summary** of 5 to 10 lines in English (for manuscripts of more than 10 pages also in German) and a short list of key-words and corresponding AMS (MOS) classification numbers are needed.

The precise address of each author with full name, academic title etc. is to be added. Information on address changes is appreciated.

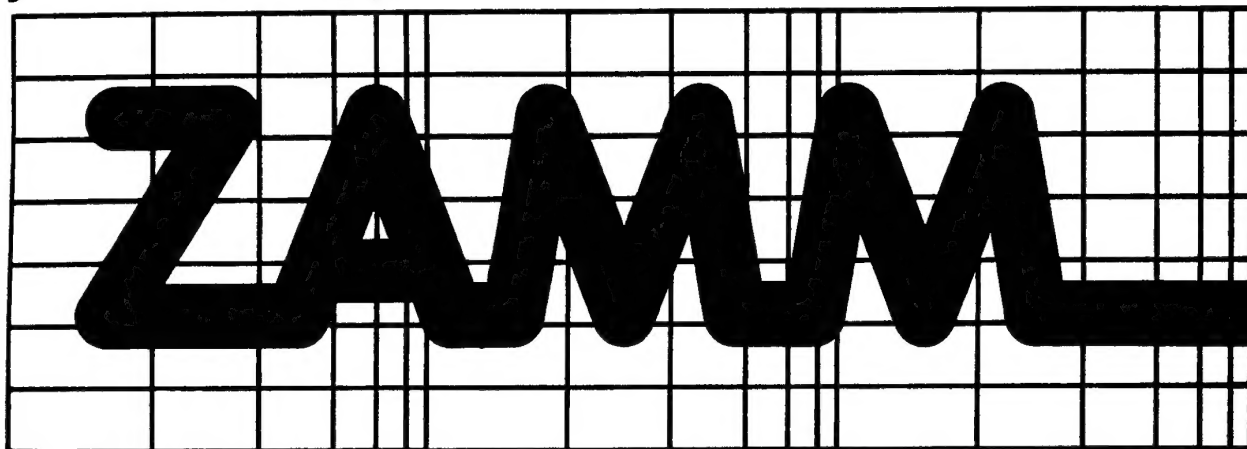
### Assignment of Rights for Journal Contributions

With the acceptance of the manuscript for publication in ZAMM – Zeitschrift für Angewandte Mathematik und Mechanik, WILEY-VCH Verlag GmbH acquires exclusively for three years from the date on which the article is published and thereafter for the full term of copyright, including any future extensions, on a nonexclusive basis all publishing rights including those of pre-publication, reprinting, translating, other forms of reproduction e.g. by photocopy, microform, or other means including machine-readable forms like CD-ROM, CD-I, DVD, diskettes, electronic storage and publishing via Local Area Networks, Intranet and Internet, and other data networks and other forms of distribution e.g. by Document Delivery-Services of this article world-wide. This includes the right of WILEY-VCH Verlag GmbH to transfer to third parties the partial or full rights. Moreover, the provisions of laws of the Federal Republic of Germany apply.

Editorial Office „ZAMM“  
University of Potsdam  
Institute of Mathematics  
POB 60 15 53  
D-14415 Potsdam  
Germany

REPORT DOCUMENTATION PAGE			Form Approved OMB No. 0704-0188	
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.				
1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE  April 2000	3. REPORT TYPE AND DATES COVERED  Conference Proceedings		
4. TITLE AND SUBTITLE  GAMM 2000 Annual Meeting, Minisymposia, University of Gottingen, Germany, April 2-7, 2000 – Supplement 1			5. FUNDING NUMBERS  F61775-00-WF017	
6. AUTHOR(S)  Conference Committee				
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)  Deutsches Zentrum für Luft- und Raumfahrt Bunsenstr. 10 Gottingen D-37073 Germany			8. Performing Organization Report Number  N/A	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)  EOARD PSC 802 Box 14 FPO 09499-0200			10. SPONSORING/MONITORING AGENCY REPORT NUMBER  CSP 00-5017	
11. SUPPLEMENTARY NOTES  Conference Proceedings in three volumes. Journal of Applied Mathematics and Mechanics (ZAMM), Volume 81 – 2001, Supplement 1, 2 and 3, ISSN 0946-8463				
12a. DISTRIBUTION/AVAILABILITY STATEMENT  Approved for public release; distribution is unlimited.			12b. DISTRIBUTION CODE  A	
ABSTRACT (Maximum 200 words)  The Final Proceedings for GAMM 2000, 2-7 April 2000. This is an interdisciplinary conference in Fluid Mechanics and Numerics in general and include topics such as Aspects of Vortex Dynamics; Controlling Chaos; Multi-scale Theory; Flow Control and optimization; Gas Dynamic; Aerothermodynamic; Turbulence; and others.				
14. SUBJECT TERMS  EOARD, Fluid mechanics, Boundary layer			15. NUMBER OF PAGES  3 Volumes	
			16. PRICE CODE	
17. SECURITY CLASSIFICATION OF REPORT  UNCLASSIFIED	18. SECURITY CLASSIFICATION OF THIS PAGE  UNCLASSIFIED	19. SECURITY CLASSIFICATION OF ABSTRACT  UNCLASSIFIED	20. LIMITATION OF ABSTRACT  UL	

# Journal of Applied Mathematics and Mechanics



Zeitschrift für Angewandte Mathematik und Mechanik

Founded by Richard von Mises in 1921

Edited in cooperation with Universität Potsdam  
and Gesellschaft für Angewandte Mathematik und Mechanik e. V. (GAMM)

Editors-in-Chief: E. Kreuzer, R. Mennicken, H. Schwetlick

## **GAMM 2000**

Annual Meeting

Minisymposia

University of Göttingen, Germany  
April 2–7, 2000

Edited by Gerd E. A. Meier, DLR Göttingen

 **WILEY-VCH**

Volume 81 · 2001  
Supplement 1

AQ F01-12-2611



and

- European Office of Aerospace Research & Development (EOARD)
- DaimlerChrysler Aerospace Airbus GmbH
- EAM Göttingen
- Sparkasse Göttingen
- Druckerei Kinzel, Göttingen
- Walter de Gruyter & Co.
- B.G. Teubner GmbH Stuttgart
- Wiley-VCH Verlag
- Springer Verlag
- *Editions Elsevier*

*We greatly appreciate the generous support given by these institutions.*

A conference of this size is the result of a collaborative effort. We would like to thank the plenary speakers, the organizers of the minisymposia and sections, the local organizers and all the authors for their contributions to the annual meeting and to these proceedings. We also thank the GAMM presidency and Mr. Trinkl from WILEY-VCH for their cooperation in the editorial process.

Special thanks are due to the external referees without whose help and expert advice a fair evaluation of the submitted manuscripts would not have been possible.

Finally we thank Elisabeth Winkels for her continuous assistance in preparing the three volumes of these proceedings.

Göttingen, March 2001

The Editor  
Prof. Dr. Gerd E. A. Meier

## Table of Contents

Preface .....	S3
<b>Minisymposium 1 – Statistics for Strongly Dependent Data</b>	
<i>Organizers:</i> M. Denker, A. Munk	
Ruymgaart, F.	
Some Results for Locally Dependent Arrays .....	S9
<b>Minisymposium 2 – Regularization Methods for Inverse Problems in Differential Equations</b>	
<i>Organizers:</i> M. Hanke-Bourgeois, O. Scherzer	
Bitterlich, S., Knabner, P.	
Formfree and Cascadic Identification of Material Laws in (Un)Saturated Fluid Flow from Column Experiments .....	S13
Hintermüller, M., Kunisch, K.	
Inverse Problems for Elastohydrodynamic Models .....	S17
<b>Minisymposium 3 – Parallel Algorithms, Domain Decomposition Methods</b>	
<i>Organizers:</i> R. H. W. Hoppe, U. Langer	
Krause, R. H., Wohlmuth, B. I.	
Domain Decomposition Methods on Nonmatching Grids and Some Applications to Linear Elasticity Problems .....	S21
Zumbusch, G.	
On the Quality of Space-filling Curve Induced Partitions .....	S25
<b>Minisymposium 4 – Non-Selfadjoint Eigenvalue Problems and Applications</b>	
<i>Organizers:</i> M. Plum, Ch. Tretter	
Denzler, J.	
Branch Classification for the Spectrum of the Orr-Sommerfeld Equation for Plane Couette Flow .....	S29
<b>Minisymposium 5 – Theory and Numerics of the Navier-Stokes Equations</b>	
<i>Organizer:</i> R. Rautmann	
Serre, E., Raspo, I., Bontoux, P., Peyret, R.	
Spectral Solution of the Navier-Stokes Equations for Rotating Flows .....	S33
Solonnikov, V. A.	
$L_p$ -estimates of Solutions of Initial-boundary Value Problem for Generalized Stokes Equations .....	S37
<b>Minisymposium 8 – Active Control of Vibration</b>	
<i>Organizers:</i> D. Guicking	
Schirmacher, R.	
The Use of Active Vibration Control for the Reduction of ICE Interior Noise .....	S41
Stöbener, U., Gaul, L.	
Active Vibration and Noise Control by Hybrid Active Acoustic Panels .....	S45
<b>Minisymposium 9 – Friction Induced Vibrations in Traction and Braking Systems</b>	
<i>Organizers:</i> K. Knothe, W. Wedig	
Allgaier, R., Gaul, L., Keiper, W., Willner, K.	
Friction Induced Vibrations and Mode Lock-In .....	S49

Ams, A., Schmalfuß, C., Wedig, W. Experimentelle und theoretische Untersuchungen an Scheibenbremsen .....	S53
Ertz, M., Knothe, K. Einfluss von Temperatur und Rauheit auf den Kraftschluss zwischen Rad und Schiene .....	S57
Müller, S., Kögel, R. Numerical Simulation of Roll-Slip Oscillations in Locomotive Drives .....	S61
Rudolph, M., Popp, K. Modellbildung reibungsselbsterregter Bremsenschwingungen .....	S65
Schmieg, H., Vielsack, P. Selbsterregte Reibschwingungen bei konstant verzögertem Vortrieb am Beispiel einer Scheibenbremse .....	S69
 <b>Minisymposium 10</b> – Prandtl's Boundary-Layer Theory and Beyond <i>Organizers:</i> E. Krause, L. Ting	
Gersten, K. Asymptotic Theory for Turbulent Shear Flows at High Reynolds Numbers .....	S73
Kluwick, A. Laminar Boundary Layer Separation, Fast and Slow .....	S77
Margerit, D., Barkley, D. Singular Perturbation Equations for 3-D Excitable Media .....	S81
Ting, L., Van Dyke, M. Prandtl's Boundary Layer Theory, before and to Matched Asymptotics .....	S85
 <b>Minisymposium 12</b> – Geomechanics of Tunnelling <i>Organizer:</i> H. A. Mang	
Macht, J., Lackner, R., Hellmich, Ch., Mang, H. A. Geomechanics of Tunneling in Squeezing Rock .....	S89
Oettl, G., Stark, R. F., Hofstetter, G. Verification of a Fully Coupled FE Model for Tunnelling under Compressed Air .....	S93
Schweiger, H. F., Schuller, H. Multilaminate Models for Numerical Simulation of Tunnel Excavations According to NATM .....	S97
Sterpi, D. Strain Localisation Effects in the Stability Analysis of Shallow Openings .....	S101
 <b>Minisymposium 13</b> – Identification of Non-Linear Systems <i>Organizers:</i> H. G. Natke, K. Popp	
Bruns, J.-U., Popp, K. Nonlinearity Tests .....	S105
Grotjahn, M., Heimann, B. Identification of Industrial Robots' Nonlinear Dynamics .....	S109
Merkwirth, C., Parlitz, U. Modeling Chaotic and Spatially Extended Systems .....	S113
Müller, P. C. Nonlinearity Estimation by PI-Observers: Theory and Applications .....	S117
Tomlinson, G. R. Frequency Response Functions: Validity and Usefulness in Nonlinear System Identification .....	S119
Unbehauen, H. Identifikation nichtlinearer Systeme aus regelungstechnischer Sicht .....	S121

# **Minisymposium 14** – Plasticity and Metal Forming Processes – On the Significance of the Prandtl-Reuß Approach

*Organizer:* D. Besdo

Böhlke, T., Bertram, A.

The 4th-Order Isotropic Tensor Function of a Symmetric 2nd-Order Tensor with Applications to Anisotropic Elasto-Plasticity ..... S125

Pawelski, H.

Theorie und Praxis des Dressierwalzens unter Berücksichtigung der Oberflächenveränderung ..... S129

# **Minisymposium 15** – Plasticity and Ductile Fracture

*Organizers:* F. D. Fischer, D. Gross

Brocks, W., Besson, J., Chabanet, O., Scheider, I., Steglich, D.

Modeling of Crack Growth in Sheet Metal ..... S133

Gumbsch, P.

Atomistische Aspekte des Bruchs ..... S137

# **Minisymposium 16** – Vortex Systems and Farfield Sound

*Organizers:* E. Krause, C. Tung

Agarwal, R. K.

Acoustic Radiation due to Gust-Airfoil and Blade-Vortex Interactions ..... S141

Blackmore, D., Knio, O.

Hamiltonian Structure for Vortex Filament Flows ..... S145

Saito, S.

Vortex Capturing and Noise Prediction by Moving Overlapped Grid Method ..... S149

Wagner, S.

Calculation of Far Field Sound due to Blade-Vortex Interaction on a Helicopter Rotor ..... S151

# **Minisymposium 17** – Compressible Multiphase Flows

*Organizer:* M. Rein

Bilicki, Z.

Wave Phenomena in Bubbly Liquid-Vapour Mixture ..... S155

Friedel, L.

Fluidynamic Critical Two-Phase (Gas/Liquid) Flow State in Non Ideal Flow Geometries

**Author Index** ..... S161

# Some Results For Locally Dependent Arrays

Frits H. Ruymgaart  
 Department of Mathematics and Statistics  
 Texas Tech University  
 Lubbock, TX 79414  
 USA

## 1. Preliminaries

A locally dependent array (l.d.a.) of order  $m = m(n)$  is a triangular array of real-valued random variables  $\{X_{n,1}, \dots, X_{n,n}, n \in \mathbb{N}\}$  such that the variables in the  $n$ -th row are  $(m-1)$ -dependent,  $1 \leq m \leq n$ . In order to have sufficiently many independent components in a row it will always be assumed that  $m/n \rightarrow 0$ , as  $n \rightarrow \infty$ . Such arrays have been studied in the literature as objects of interest in their own right (Berk (1973), Barbour (1990), Reinert (1995, 1996)), and as a tool to approximate certain time series (Portnoy (1991), Chanda and Ruymgaart (1990, 1991)). Time series with Volterra expansions seem particularly suited for such approximations.

Here we want to focus on properties of the empirical process of the array which are useful when nonparametric curves like autoregression functions are to be estimated (Chanda and Ruymgaart (1991)). Some results from Einmahl and Ruymgaart (1998) including weak convergence in  $D([0,1])$  will be surveyed in Section 2. A conjecture on weak convergence in  $L^2([0,1])$  for a wider class of l.d.a.'s with application to the empirical process of a linear process is formulated in Section 3.

Having the empirical process in mind for the most part we will restrict ourselves to arrays that satisfy the following assumptions:

- all the  $X_{n,i}$  assume values in  $[0,1]$ ;
- the  $X_{n,i}$  are rowwise identically distributed with c.d.f.  $F_n$ .

The actual strength of the dependence within a row can be locally strong and is in some sense further specified by the variance, of a block. Under the present conditions we always have

$$\text{Var}\left(\sum_{j=1}^k X_{n,i+j}\right) \leq Ckm, \quad (1.1)$$

for some generic  $0 < C < \infty$ , but the l.h.s. may be actually of smaller order than the r.h.s. The following example shows that, typically, local dependence foregoes long range dependence.

**Linear Processes.** Let us approximate the linear process  $X_i = \sum_{k=-\infty}^{\infty} a_k \varepsilon_{i-k}$ , where the  $\varepsilon_j$  are i.i.d.  $(0,1)$ , with the array  $X_{n,i} = \sum_{k=-m}^m a_k \varepsilon_{i-k}$  ( $i = 1, \dots, n$ ). Let us assume that  $|a_k| \approx 1/(1+|k|^\delta)$  for some  $\delta > \frac{1}{2}$ . For useful approximation we need on the one hand that

$$\max_{1 \leq i \leq n} |X_i - X_{n,i}| = o_p(n^{-\varepsilon}), \text{ for some } \varepsilon > 0, \quad (1.2)$$

and on the other hand  $m \ll n$ . Straightforward calculation shows that this can be only achieved for  $\delta > 1 + \varepsilon$ . For such  $\delta$  the linear process is *not* long range dependent (Beran (1994)), and the variance in (1.1) is bounded by  $Ck$



and consequently is of the same order as in the i.i.d. case. For long range dependent moving averages the empirical process has been studied in Ho and Hsing (1996) and for certain long range dependent stationary processes by Dehling and Taqqu (1989).

## 2. Some Tools And Weak Convergence

In order to prove weak convergence of the empirical process of the l.d.a. a central limit theorem will be needed to deal with the finite dimensional distributions, and a fluctuation inequality to deal with the tightness in  $D([0, 1])$ .

**Central Limit Theorem.** In addition to the assumptions of Section 1 let us assume that for  $1 \leq \ell = \ell(n) \leq m$

$$\text{l.h.s. of (1.1)} \leq C \frac{k}{n} s_n^2, \quad s_n^2/(n\ell) \rightarrow \sigma^2 \geq 0, \quad nm^{1+\delta}/s_n^{2+\delta} \rightarrow 0 \text{ for some } \delta > 0, \quad (2.1)$$

as  $n \rightarrow \infty$ , where  $s_n^2 = \text{Var}(\sum_{i=1}^n X_{n,i})$ . Then we have

$$\frac{1}{\sqrt{n\ell}} \sum_{i=1}^n X_{n,i} \rightarrow_d N(0, \sigma^2), \text{ as } n \rightarrow \infty. \quad (2.2)$$

**Fluctuation Theorem.** Let us write

$$\hat{F}_n(t) = \frac{1}{n} \sum_{i=1}^n 1_{[0,t]}(X_{n,i}), \quad 0 \leq t \leq 1, \quad (2.3)$$

$$\Delta_n = \hat{F}_n - F_n, \quad (2.4)$$

and, for any interval  $I = (a, b]$  let  $F_n\{I\} = F_n(b) - F_n(a)$ ,  $\hat{F}_n\{I\} = \hat{F}_n(b) - \hat{F}_n(a)$ ,  $\Delta_n\{I\} = \hat{F}_n\{I\} - F_n\{I\}$ . Fix an interval  $I_0$  with  $F_n\{I_0\} \leq \frac{1}{2}$ . Then there exists  $0 < \varepsilon < 1$  such that for each  $n$  and  $\lambda \geq 0$

$$\mathbb{P}\{\sup_{I \subset I_0} |\Delta_n\{I\}| \geq \lambda\} \leq C(\varepsilon) \exp\left(\frac{-(1-\varepsilon)n\lambda^2}{2mF_n\{I_0\}} \psi\left(\frac{\sqrt{n}\lambda}{\sqrt{m\lfloor \frac{n}{m} \rfloor} F_n\{I_0\}}\right)\right), \quad (2.5)$$

where  $\psi(x) = 2x^{-2}\{(1+x)\log(1+x) - x\}$ ,  $x > 0$ , and  $\psi(0) = 1$ .

In order to be able to derive tightness in  $D([0, 1])$  from (2.5) it is necessary that the proper scaling factor of  $\Delta_n$  is  $\sqrt{n/m}$  so that the factor  $m$  in the denominator of the exponent will cancel. This places a technical restriction on which we will further comment in Section 3.

**Weak Convergence in  $D([0, 1])$ .** Suppose we have

$$\sup_{0 \leq t \leq 1} |F_n(t) - F(t)| \rightarrow 0, \text{ as } n \rightarrow \infty, \quad (2.6)$$

for some c.d.f.  $F$ , and

$$\frac{n}{m} \text{Cov}(\Delta_n(s), \Delta_n(t)) \rightarrow \Gamma(s, t), \text{ as } n \rightarrow \infty, \quad (2.7)$$

for some nondegenerate covariance function  $\Gamma$  on the unit square. Then there exists a Gaussian process  $\mathcal{G}$  in the space  $D([0, 1])$  endowed with the  $J_1$ -topology such that

$$\sqrt{\frac{n}{m}} \Delta_n \rightarrow_d \mathcal{G}, \text{ as } n \rightarrow \infty, \text{ in } D([0, 1]). \quad (2.8)$$

The process  $\mathcal{G}$  is centered and has covariance function  $\Gamma$ .

### 3. Weak Convergence In A Hilbert Space

Unfortunately it can be seen from the results in the example of Section 1 that for linear processes (with  $\delta > 1$ ) condition (2.7) will not be satisfied. In fact the convergence rate is much faster. Although our proof of weak convergence in  $D([0, 1])$  breaks down we may try to prove weak convergence in  $L^2([0, 1])$ ; see Laha and Rohatgi (1979).

Let us consider an l.d.a. with strictly stationary rows, and assume that (cf. (2.1), (2.7))

$$\text{Var}\left(\sum_{i=1}^k 1_{[0, t]}(X_{n, i})\right) \leq C k \ell, \quad (3.1)$$

for all  $t$ , and

$$\frac{n}{\ell} \text{Cov}(\Delta_n(s), \Delta_n(t)) \rightarrow \Gamma(s, t), \text{ as } n \rightarrow \infty, \quad (3.2)$$

in addition to (2.6). In much the same way as in the proof of the CLT the process  $\Delta_n$  can be written as the sum of i.i.d. processes and a remainder. Further conditions to ensure the weak convergence

$$\sqrt{\frac{n}{\ell}} \Delta_n \rightarrow_d \mathcal{G}, \text{ as } n \rightarrow \infty, \text{ in } L^2([0, 1]), \quad (3.3)$$

should be mild. Of course such a result holds true for random variables with values in any bounded interval  $[a, b]$ .

Returning to the linear process let us assume for convenience that the  $X_i$  assume values in  $[-1, 1]$ , and let  $F$  be their common c.d.f. Denote their empirical process by  $\Delta_n^*$ , based on the empirical c.d.f.  $\hat{F}_n^*$ , and let  $\Delta_n$  be the empirical process of the l.d.a.  $X_{n,1}, \dots, X_{n,n}$  (for suitable  $m$ ), based on the empirical c.d.f.  $\hat{F}_n$ . If  $\delta > 1$  for this array (3.2) will hold true with  $\ell = 1$ . Since

$$n \int_{-1}^1 \{\hat{F}_n^*(t) - \hat{F}_n(t)\}^2 dt \leq n \max_{1 \leq i \leq n} |X_i - X_{n,i}|, \quad (3.4)$$

the truncation index  $m = m(n)$  can be determined in such a way that the l.h.s. of (3.4) tends in probability to 0, as  $n \rightarrow \infty$ , provided that  $\delta > 3\frac{1}{2}$ . If we assume that  $n \int_{-1}^1 \{F(t) - F_n(t)\}^2 dt \rightarrow 0$ , as  $n \rightarrow \infty$ , the weak convergence (3.3) for the empirical process of the l.d.a. (with  $\ell = 1$ ) entails the weak convergence

$$\sqrt{n}\Delta_n^* \rightarrow_d \mathcal{G}, \text{ as } n \rightarrow \infty, \text{ in } L^2([0, 1]), \quad (3.5)$$

for the empirical process of the linear process.

### Acknowledgements

The authors are grateful to Manfred Denker and Axel Munk for their invitation to present this paper at the GAMM 2000 minisymposium "Statistics for Strongly Dependent Data".

### 4. References

- 1 Barbour, A. D. (1990). Stein's method for diffusion approximations. *Probab. Th. Rel. Fields* **84**, 297-322.
- 2 Beran, J. (1994). *Statistics for Long-Memory Processes*. Chapman and Hall, New York.
- 3 Berk, K. N. (1973). A central limit theorem for  $m$ -dependent random variables with unbounded  $m$ . *Ann. Probab.* **1**, 352-354.
- 4 Chanda, K. C. and Ruymgaart, F. H. (1990). General linear processes: a property of the empirical process applied to density and mode estimation. *J. Time Ser. Anal.* **11**, 185-199.
- 5 Chanda, K. C. and Ruymgaart, F. H. (1991). Curve estimation for  $m_n$ -decomposable time series including bilinear processes. *J. Multivar. Anal.* **38**, 149-166.
- 6 Dehling, H. and Taqqu, M. S. (1989). The empirical process of some long-range dependent sequences with application to  $U$ -statistics. *Ann. Statist.* **17**, 1767-1783.
- 7 Einmahl, J. H. J. and Ruymgaart, F. H. (1998). Some results for empirical processes of locally dependent time series. Report COSOR 98-01, Eindhoven Univ. Tech.
- 8 Ho, H.-C. and Hsing, T. (1996). On the asymptotic expansion of the empirical process of long-memory moving averages. *Ann. Statist.* **24**, 992-1024.
- 9 Laha, R. G. and Rohatgi, V. K. (1979). *Probability Theory*. Wiley, New York.
- 10 Portnoy, S. (1991). Asymptotic behavior of regression quantiles in nonstationary dependent cases. *J. Multivar. Anal.* **38**, 100-113.
- 11 Reinert, G. (1995). A weak law of large numbers for empirical measures via Stein's method. *Ann. Probab.* **23**, 334-354.
- 12 Reinert, G. (1996). Gaussian approximations for random measures via Stein's method. Preprint.

BITTERLICH, S., KNABNER, P.

## Formfree and Cascadic Identification of Material Laws in (Un)Saturated Fluid Flow from Column Experiments

*Simulations of processes in porous media often are based on mathematical models in form of partial differential equations or systems of such equations. The processes in porous media critically depend on knowledge of material laws, characteristic material properties of the considered porous media, which are represented in the model equations by coefficient functions. Therefore, the material properties have to be determined by suitable experiments. Usually such experiments can only reveal an output of the system caused by the material properties to a certain input. The material property itself has to be identified from this by mathematical and numerical methods (inverse modeling). This paper presents an approach for formfree and cascadic identification of nonlinear coefficient functions in a partial differential equation based on output least squares minimization.*

### 1. (Un)Saturated Fluid Flow in Porous Media

Flow in an (un)saturated porous medium is described by the well-known Richards equation, which in its pressure head form is given by a mass balance equation

$$\partial_t \Theta(\psi) + \nabla \cdot \vec{q} = 0$$

and Darcy's law

$$\vec{q} = -K(\psi) \nabla(\psi + z),$$

where  $\psi = \psi(\vec{x}, t)$  [Length] is the pressure head,  $\vec{q} = \vec{q}(\vec{x}, t)$  [Length/Time] is the volumetric flow rate per unit surface area and  $z$  is the height against gravitational direction.

The coefficients in above equations characterize the hydraulic properties of the medium, fluid retention  $\Theta(\psi)$  [-] and hydraulic conductivity  $K(\psi)$  [Length/Time]. These functions are monotone increasing and constant in the saturated region ( $\psi \geq 0$ ). Soil scientists work mostly with the van Genuchten-Mualem model (see [6]). Soil column outflow experiments are a suitable method to determine the hydraulic functions (see [2]).

Column experiments exhibit a flow regime only in one direction such that a spatially one-dimensional model suffices. The experiments involve draining a vertically oriented soil column of length  $L$  with known initial pressure head distribution  $\psi_0(x)$  near saturation by slowly decreasing the pressure head  $h(t)$  at the lower boundary. Mathematically this is modeled by a Dirichlet boundary condition. The flux at the upper boundary is adjusted to  $q = 0$ , therefore a homogeneous flux boundary condition is used. The flux  $f(t)$  is measured at the outlet at the lower boundary:

$$f(t) = q(L, t).$$

Further the pressure head  $g(t)$  is measured at the upper boundary:

$$g(t) = \psi(0, t).$$

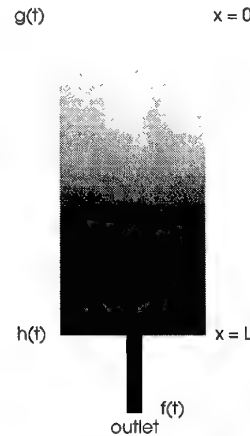


Figure 1: Soil column

The physical properties of the experiments allow us to assume that it suffices to describe the flow in the column by the Richards equation in one spatial dimension:

$$\left. \begin{aligned} \partial_t \Theta(\psi) - \partial_x (K(\psi)(\partial_x \psi - 1)) &= 0, & (x, t) &\in (0, L) \times (0, T) \\ \psi(x, 0) &= \psi_0(x), & x &\in (0, L) \\ K(\psi(0, t))(\partial_x \psi - 1) &= 0, & \psi(L, t) &= h(t), \quad t \in (0, T). \end{aligned} \right\} \quad (1)$$

Solving the model equation (1) for given hydraulic functions  $\Theta$  and  $K$  and assigning the measurements  $f(t)$  and  $g(t)$  to  $\Theta$  and  $K$  characterize the direct problem (DP:  $(\Theta, K) \mapsto (f, g)$ ). The inverse problem consists of determining the hydraulic functions  $\Theta$  and  $K$  from given measurements  $f(t)$  and  $g(t)$  (IP:  $(f, g) \mapsto (\Theta, K)$ ). In general, inverse problems are ill-posed, therefore we need an appropriate regularization strategy for a stable solution process of the identification problem.

## 2. Identifiability

Assume  $f$  and  $g$  are obtained by exact, error-free measurements. If the mapping IP is injective, the coefficients  $\Theta$  and  $K$  can be uniquely identified. To prove this property of the inverse problem we apply the method of integral identities (see [1],[3]).

According to [4], we define distinguishability of real continuous functions.

**Definition 1.** Two real, continuous functions  $\alpha, \beta : [a, b] \rightarrow \mathbb{R}$  are called *distinguishable* on  $[a, b]$ , if there exists a finite partition  $a = x_1 < \dots < x_n = b$ , such that on every interval  $(x_{i-1}, x_i)$  either  $\alpha(x) < \beta(x)$  or  $\alpha(x) > \beta(x)$  or  $\alpha(x) = \beta(x)$  for  $x \in (x_{i-1}, x_i)$  is valid. One of the cases " $<$ " and " $>$ " has to be valid at least once.

Now, we can present the identifiability theorem.

**Theorem 2.** Let  $(f_1, g_1)$  and  $(f_2, g_2)$  the solution of the direct problem for sufficiently smooth coefficients  $(\Theta_1, K_1)$  and  $(\Theta_2, K_2)$ , respectively. If the applied suction  $h(t)$  at the lower boundary is smooth and monotone decreasing in time and the pairs  $(\Theta_1, K_1)$  and  $(\Theta_2, K_2)$  are distinguishable on the interval  $[\psi^*, 0]$  for  $\psi^* = -L + h(T)$  (the lower boundary of the experimentally considered pressure head region) with  $\Theta_1(0) = \Theta_2(0)$ , then  $(f_1, g_1)$  and  $(f_2, g_2)$  are not identical on  $[0, T]$ .

**Proof.** See [1].

In the sense of the above theorem it is meaningful to assume that the inverse problem is uniquely solvable. Consequently the hydraulic functions fluid retention  $\Theta$  (if  $\Theta(0)$  is known) and hydraulic conductivity  $K$  are identifiable from soil column outflow experiments.



Figure 2: Hydraulic properties (dashed) and identified hydraulic properties (piecewise quadratic) for 9 degrees of freedom (solid), left: fluid retention  $\Theta$ , right: hydraulic conductivity  $K$ , "optimal".

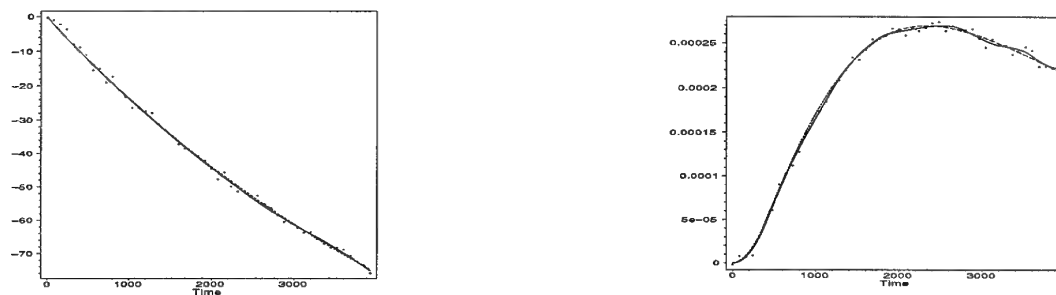


Figure 3: Pressure head (left) at the upper boundary and flux (right) at the outlet (dashed: original, solid: reconstructed, points: "measured"), "optimal".



### 3. Stabilization of the Inverse Problem

The inverse problem is stabilized by a parametrization of the hydraulic functions. We are looking for an unbiased parametrization, which does not take other a-priori shape informations as basic physical properties like e.g. monotonicity into account.

A general approach uses spline functions to parameterize the functions that have to be identified. In this way we obtain e.g. a piecewise linear or a piecewise quadratic parametrization. Then the unknown functions are defined by real parameter vectors  $p^\Theta$  and  $p^K$  of a finite dimension  $r$ , called number of degrees of freedom. A special aspect of such a parametrization is the fact that the fluid retention and the hydraulic conductivity are not coupled like in the van Genuchten-Mualem model. The low smoothness of piecewise linear functions  $\Theta$  and  $K$  leads to low smoothness of the observations  $f(t)$  and  $g(t)$  corresponding to the measurements. Therefore an approach with quadratic B-splines is profitable. Apart from a local basis we can use a concept of hierarchical and multi scale basis. Such parametrizations correspond to projection methods for the regularization of inverse problems. Of course, a threshold  $r_{max}$  (depending on the discretization error, measurement error and the type of parametrization) exists, such that a parametrization with more than  $r_{max}$  degrees of freedom entails instabilities. Note that due to the local dependence of the unknown functions on the parameters only a reconstruction for the range of  $\psi$ -values, which is covered by the experiment, can be expected.



Figure 4: Hydraulic properties (dashed) and identified hydraulic properties (piecewise quadratic) for 9 degrees of freedom (solid), left: fluid retention  $\Theta$ , right: hydraulic conductivity  $K$ , "non-optimal".



Figure 5: Pressure head (left) at the upper boundary and flux (right) at the outlet (dashed: original, solid: reconstructed, points: "measured"), "non-optimal".

### 4. Numerical Identification by Output Least Squares

The identification problem is transformed into a minimization problem for an error functional. We deal with usual output least squares functionals. The discretization of the coefficient functions leads to a finite dimensional optimization problem, where parameter vectors  $p^\Theta$  and  $p^K$  have to be found which minimize

$$J(p^\Theta, p^K) = \sum_{i=1}^M \alpha_i (f(\hat{t}^i) - \hat{f}^i)^2 + \sum_{i=1}^N \beta_i (g(\tilde{t}^i) - \tilde{g}^i)^2 \quad (2)$$

with positive weighting factors  $\alpha_i$ ,  $\beta_i$  and measured values  $\hat{f}^i$  and  $\tilde{g}^i$  at time points  $\hat{t}^i$  and  $\tilde{t}^i$ , respectively. For a numerical identification procedure the direct problem has to be discretized, too. We use a hybrid mixed finite element method described in [5].

Efficient optimization algorithms need the value of the error functional (2) and its gradient. The gradient of the error functional can be evaluated in two ways:

- a) Finite difference method: We have to solve the direct problem  $2r$  times for the one-sided and  $4r$  times for the central difference quotient to compute an approximation of the gradient.
- b) Adjoint method: Differentiation of the discrete problem leads to an adjoint system of equations for the gradient. The combined solution process needs here about twice the CPU time of solving the direct problem.

Adding a penalty term to the error functional (in accordance with the Tikhonov regularization) and constraints (e.g. monotonicity conditions) further stabilize the minimization procedure. Under the assumption that the Mualem-relation

$$K^*(\Theta) = K_s \Theta^{1/2} \left[ \frac{\int_0^\Theta [1/\psi(s)] ds}{\int_0^1 [1/\psi(s)] ds} \right]^2 \quad (3)$$

where  $K_s$  is the saturated conductivity, is a realistic model of hydraulic conductivity, an appropriate penalty term is given by deviations between the hydraulic conductivity and the Mualem-model

$$\alpha \sum_i |K(\Theta_i) - K^*(\Theta_i)|^2 \quad (4)$$

for a small positive regularization parameter  $\alpha$ . The values  $K^*(\Theta_i)$  in the penalty term (4) can be obtained by a quadrature rule. For it we approximate the retention function  $\Theta$  exponentially for values of the pressure head less than  $\psi^*$ .

Major problems of minimizing the highly nonlinear functional are the high sensitivity to the initial value and the slow convergence. To eliminate these problems we embed the identification procedure in a multi level algorithm: start minimization with the least possible number of degrees of freedom, interpolate the result of the optimization for a parametrization with one or more added degrees of freedom and use this as next initial value. In this context the hierarchical concept generates a scale by scale optimization. The values of the hydraulic functions at saturation  $\Theta_s$  and  $K_s$  usually are known from independent experiments and can be used to determine initial values for a parametrization with the least possible number of degrees of freedom.

We simulate an experiment for a column with a van Genuchten-Mualem parametrization of the hydraulic functions and disturb the "simulated measurements" by a gaussian distributed noise (5%). If a local basis is used for the parametrization, then the identification results depend on the so called refinement strategy for the increase of the degrees of freedom during the multi level algorithm. The results for two different refinement strategies are shown in figures 2 and 3 ("optimal") and in figures 4 and 5 ("non-optimal").

## 5. References

- 1 DUCHATEAU, P.: An inverse problem for hydraulic properties of porous media. SIAM J. Math. Anal., Vol. 28, No. 3, 1997, 611-632.
- 2 DURNER, W., SCHULTZE, B., ZURMÜHL, T.: State-of-Art in Inverse Modeling of Inflow/Outflow Experiments. Proc. Int. Workshop on Characterization and Measurement of the Hydraulic Properties of Unsaturated Porous Media, University of California, Riverside, CA., 1999, 661-681.
- 3 IGLER, B.A.: Identification of Nonlinear Coefficient Functions in Reactive Transport through Porous Media. PhD thesis, Friedrich-Alexander-Universität Erlangen-Nürnberg, 1998.
- 4 IGLER, B.A., KNABNER, P.: Structural Identification of Nonlinear Coefficient Functions in Transport Processes Through Porous Media. in: Lectures on Applied Mathematics, Springer Verlag, 2000, 157-178.
- 5 KNABNER, P., SCHNEID, E.: Numerical Solution of Unsteady Saturated/Unsaturated Flow Through Porous Media. Proceedings of the 2nd Summer Conference Numerical Modelling in Continuum Mechanics, Prag, 1997.
- 6 VAN GENUCHTEN, M.TH.: A closed-form equation for predicting the hydraulic conductivity of unsaturated soils. Soil Sci. Soc. Am. J., Vol. 44, 1980, 892-898.

*Addresses:* SANDRO BITTERLICH, PROF. DR. PETER KNABNER, Institute for Applied Mathematics, Universität Erlangen, Martensstrasse 3, D 91058 Erlangen, Germany

M. HINTERMÜLLER AND K. KUNISCH

## Inverse Problems for Elastohydrodynamic Models

*Inverse coefficient problems for variational inequalities arising in the elastohydrodynamic lubrication of a journal bearing are presented. The height of the gap between two rotating surfaces denotes the distributed parameter that has to be identified from estimates of the pressure in the lubricant between the surfaces. The variational inequality approach which includes the phenomenon of cavitation, i.e. the situation where the gap is not entirely filled by the lubricant, reduces to the Reynolds lubrication equation under fully-flooded conditions. Utilizing a regularized least-squares formulation the problem of existence of multipliers, and the importance and derivation of a first order characterization amenable for (structured) numerical realization are addressed.*

### 1. Introduction

**The direct problem.** In this paper we consider lubrication problems for journal bearings. The schematic cross-section of a typical device is shown in Figure 1. By  $u$  we denote the height of the gap between the ball and the wall

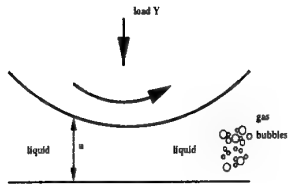


Fig. 1: Ball bearing (schematic cross-section)

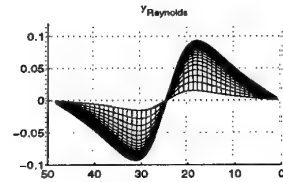


Fig. 2: Solution of Reynolds equation

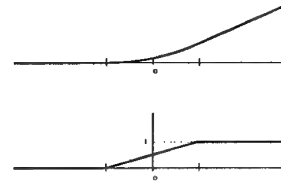


Fig. 3:  $\max_c$  (upper graph) and  $\nabla \max_c$  (lower graph).

of the bearing. We assume that this height is small compared to the other dimensions. To avoid damage, a liquid (lubricant) is used to maintain strict positivity of the gap. The ball imposes a certain load which induces a pressure  $y$  in the lubricant. It is well known that the Reynolds equation

$$-\operatorname{div}((\rho u^3/6\mu)\nabla y) = \rho s \frac{\partial u}{\partial x_2} \quad (1)$$

models the pressure distribution  $y$ . In (1)  $\rho$  and  $\mu$  denote specific densities and viscosities, respectively, which are both assumed to be constant. Moreover,  $s$  is a constant relative velocity of the two surfaces in motion. Throughout we further assume that the ambient pressure  $p_a$  is zero. Note that the weak form of (1) together with boundary conditions is the necessary and sufficient first order condition for

$$\text{minimize } \int_{\Omega} e(u)|\nabla y|^2 dx - \int_{\Omega} (Fu)y dx =: J^u(y) \quad \text{over } y \in H^1_0(\Omega), \quad (2)$$

with  $e(u) = \rho u^3/6\mu$  and  $Fu = \rho s \frac{\partial u}{\partial x_2}$ .

Unfortunately, the Reynolds equation is not always adequate. Especially, in the cases of gaseous and vapor cavitation (indicated by the bubbles in Figure 1) a pressure distribution below the ambient pressure (in our case a negative pressure distribution) is predicted; see Figure 2, where the horizontal axis displays the number of grid points in  $(0, 1)$ . This behavior does not conform to physics. In certain cases, a remedy is given by adding the constraint  $y \geq 0$  in (2), i.e. now the minimization is taken over  $K := \{y \in H^1_0(\Omega) | y \geq 0\}$ . The corresponding first order condition for the unique solution is the variational inequality

$$(e(u)\nabla y, \nabla(v - y)) - (Fu, v - y) \geq 0 \quad \text{for all } v \in K. \quad (3)$$

Above  $(\cdot, \cdot)$  denotes the usual  $L^2(\Omega)$ -inner product.

**The inverse problem.** For several reasons the problem of estimating the height of the gap  $u$  from estimates  $y_d$  of the pressure distribution  $y$  is of importance. (i) For a certain regime, from engineering experience a good estimate

of the pressure distribution may be available. Then the investigation of the possibility of damage of the bearing, i.e.  $u \leq 0$ , is of central importance. (ii) In cases of high pressure phenomena such as vapor cavitation a second equation modeling the influence of the pressure on the gap has to be taken into account. Hertzian contact theory yields

$$u = u_0 + \int_{\Omega} K(x_1, x_2) y(x_2) dx_2, \quad K(x_1, x_2) = \frac{c_0}{\sqrt{(x_1 - \hat{x}_1)^2 + (x_2 - \hat{x}_2)^2}} \quad (\text{point contact}), \quad (4)$$

where  $u_0$  denotes a constant clearance. The numerical procedure for computing  $y$  from the coupled system (3) and (4) by

(i) given  $y^n$  compute  $u^n$  from (4); (ii) solve (3) to obtain  $y^{n+1}$  and go to (i)

frequently fails to converge or converges at a low rate. A possible remedy is to reverse the procedure, i.e.

(i)<sub>r</sub> given  $y^n$  compute  $u^n$  from (3); (ii)<sub>r</sub> solve (4) to obtain  $y^{n+1}$  and go to (i)<sub>r</sub>.

Note that (i)<sub>r</sub> amounts to a parameter identification problem in a variational inequality which – compared to (ii)<sub>r</sub> – is the difficult task. Therefore, subsequently we concentrate on the inverse coefficient problem in (i)<sub>r</sub>.

## 2. First order necessary characterization

There are several approaches to identify  $u$  from estimates  $y_d$  of  $y$ . Here we concentrate on regularized least squares techniques, i.e. we consider

$$(P) \quad \begin{aligned} &\text{minimize} && \frac{1}{2} \|y(u) - y_d\|_{L^2(\Omega)}^2 + \frac{\alpha}{2} \|u\|_{H^1(\Omega)}^2 \\ &\text{s.t.} && u \in U, \\ &&& y(u) = \operatorname{argmin}\{J^u(y) | y \in K\}, \end{aligned} \quad (5)$$

where  $U = \{u \in H^1(\Omega) \cap L^\infty(\Omega) | 0 < \underline{\varepsilon} \leq u \leq \bar{\varepsilon} < +\infty\}$ , and  $\|\cdot\|_{L^2(\Omega)}$  and  $\|\cdot\|_{H^1(\Omega)}$  denote norms in  $L^2(\Omega)$  and  $H^1(\Omega)$ , respectively. Moreover,  $\alpha > 0$  is the regularization parameter.

Problem (P) is a *bilevel* optimization problem due to the fact that (5) is a minimization problem with  $u$  as a parameter. Usually solutions to problems of type (P) are characterized by a first order system involving (Lagrange) multipliers. However, here the lower level problem prevents the immediate application of the classical theory, since by replacing (5) by its first order optimality condition (3) or equivalently by

$$\begin{aligned} &(e(u) \nabla y(u), \nabla v) - (\lambda, v) = (Fu, v) \quad \text{for all } v \in H_0^1(\Omega), \\ &\lambda \geq 0, \quad y(u) \geq 0, \quad (y(u), \lambda) = 0, \end{aligned}$$

certain regularity requirements may get lost and existence of multipliers cannot be guaranteed (even in practically important situations). Therefore, our aim is to derive multipliers that exist in very general situations, and such that the corresponding first order characterization is amenable for numerical realization. Combining these two aspects contrasts most of the available literature e.g. [1] and the references in [2].

**Primal-dual penalization.** The first step towards the existence proof is the reformulation of the lower level problem (5) by means of a primal-dual penalization technique. In fact, we consider

$$\text{minimize} \quad J^u(y) + \frac{1}{2c} \|\max\{\bar{\lambda} - cy, 0\}\|_{L^2(\Omega)}^2 =: J_c^u(y) \quad \text{over } y \in H_0^1(\Omega), \quad (6)$$

where  $\bar{\lambda} \in L^2(\Omega)$ ,  $\bar{\lambda} \geq 0$ , is arbitrarily fixed, and  $c > 0$ . Note that we have replaced the explicit constraint  $y \geq 0$  by an implicit one. The first order condition for the optimal solution  $y_c(u)$  of (6) is given by

$$(e(u) \nabla y_c(u), \nabla v) - (Fu, v) - (\max\{\bar{\lambda} - cy_c(u), 0\}, v) = 0 \quad \text{for all } v \in H_0^1(\Omega). \quad (7)$$

**Lemma 1.** (i) Let  $\{c_n\} \subset \mathbb{R}^+$  be a sequence satisfying  $c_n \rightarrow \infty$  for  $n \rightarrow \infty$ . Then  $\{y_{c_n}(u)\}$  converges to  $y(u)$  (cf. (5)) strongly in  $H_0^1(\Omega)$ . (ii) The mappings  $\Phi : u \mapsto y(u)$  and  $\Phi_c : u \mapsto y_c(u)$  are completely continuous.

For the proof we refer to [2]. The aim of Lemma 1 is twofold: Due to (ii) it can be seen that (P) and the bilevel problem after reformulation of the lower level problem admit optimal solutions by using techniques similar to those in [1,2,3]. Assertion (i) can also be seen as a consistency result, i.e. by solving (6) repeatedly with increasing  $c$ , the solution to the original problem is approached.

The next theorem is a key result (for its proof see [2]): It clarifies the role of  $\bar{\lambda}$ , i.e. a suitable choice yields  $y_c(u) \geq 0$ . Moreover, employing the *concept of complementarity functions* it aims at numerical amenability. For the latter purpose the following definition is necessary.

**Definition 2.** A function  $\Theta: \mathbb{R}^2 \rightarrow \mathbb{R}$  is called *complementarity function* iff

$$\Theta(a, b) = 0 \iff a \geq 0, b \geq 0, ab = 0.$$

**Theorem 3.** Let  $\{c_n\} \subset \mathbb{R}^+$  with  $c_n \rightarrow \infty$ , and let  $\Theta$  be a complementarity function. Then: (i) for  $\bar{\lambda} := \max\{-(Fu), 0\}$  we have  $y_{c_n}(u) \geq 0$ ; (ii)  $\{\max\{\bar{\lambda} - c_n y_{c_n}(u), 0\}\}$  converges weakly in  $L^2(\Omega)$  to  $\lambda(u)$  satisfying

$$(e(u)\nabla y(u), \nabla v) - (Fu + \lambda(u), v) = 0 \text{ for all } v \in H_o^1(\Omega), \quad \Theta(\lambda(u), y(u)) = 0 \text{ a.e. in } \Omega.$$

**Regularization.** Due to the nondifferentiability of the max-operation in (7) we use a regularization technique as in [2,3], i.e. we set

$$\max_c\{x, 0\} = \begin{cases} x & \text{for } x \geq \frac{1}{2c} \\ \frac{c}{2}(x + \frac{1}{2c})^2 & \text{for } |x| \leq \frac{1}{2c} \\ 0 & \text{for } x \leq -\frac{1}{2c} \end{cases}$$

For the graphs of  $\max_c$  and its first derivative, which we denote by  $\text{sgn}_c$ , see Figure 3.

For the regularized primal-dual reformulation of (P) which is equivalent to

$$\begin{aligned} (\tilde{P}_c) \quad & \text{minimize} \quad \frac{1}{2} \|y - y_d\|_{L^2(\Omega)}^2 + \frac{\alpha}{2} \|u\|_{H^1(\Omega)}^2 \\ & \text{s.t.} \quad u \in U, \\ & (e(u)\nabla y, \nabla v) - (Fu + \max_c\{\bar{\lambda} - cy, 0\}, v) = 0 \text{ for all } v \in H_o^1(\Omega) \end{aligned}$$

the analogues of Lemma 1 and Theorem 3 hold true. Note that the problem  $(\tilde{P}_c)$  can be viewed as a *standard* optimal control problem. Thus, applying the well-known (Lagrange) multiplier technique we arrive at the necessary first order characterization (in weak form):

$$\begin{aligned} (e(\tilde{u}_c^*)\nabla \tilde{p}_c^*, \nabla v) + c(\text{sgn}_c(\bar{\lambda} - c\tilde{y}_c^*), v) + (\tilde{y}_c^* - y_d, v) &= 0 \text{ for all } v \in H_o^1(\Omega), \\ (\alpha B^* B \tilde{u}_c^* + e'(\tilde{u}_c^*)(\nabla \tilde{y}_c^* \cdot \nabla \tilde{p}_c^*) - F^* \tilde{p}_c^*, u - \tilde{u}_c^*) &\geq 0 \text{ for all } u \in U, \\ (e(\tilde{u}_c^*)\nabla \tilde{y}_c^*, \nabla v) - (F\tilde{u}_c^* + \max_c\{\bar{\lambda} - c\tilde{y}_c^*, 0\}, v) &= 0 \text{ for all } v \in H_o^1(\Omega). \end{aligned}$$

Here  $(\tilde{y}_c^*, \tilde{u}_c^*) \in H_o^1(\Omega) \times U$  is the optimal solution of  $(\tilde{P}_c)$ ,  $\tilde{p}_c^*$  denotes the associated adjoint state,  $B$  represents the embedding operator from  $H^1(\Omega)$  to  $L^2(\Omega)$ , and  $F^*$ ,  $B^*$  are the dual operators of  $F$ ,  $B$ , respectively. Moreover,  $e'(\cdot)$  is the derivative of  $e(\cdot)$ .

**Passage to the limit.** Combining the previous results and passing to the limit for  $c \rightarrow \infty$  yields

**Theorem 4.** Let  $\{c_n\} \subset \mathbb{R}^+$  with  $c_n \rightarrow \infty$ , and let  $\Theta$  be a complementarity function. Then  $(\tilde{y}_c^*, \tilde{u}_c^*)$  converges strongly-weakly in  $H_o^1(\Omega) \times U$  to  $(y^*, u^*)$ , an optimal solution to (P). Moreover, there exist multipliers  $(p^*, \mu^*) \in H_o^1(\Omega) \times (L^\infty(\Omega))^*$  satisfying

$$\begin{aligned} (e(u^*)\nabla p^*, \nabla v) + (\mu^*, v) + (y^* - y_d, v) &= 0 \text{ for all } v \in H_o^1(\Omega), \\ (\alpha B^* B u^* + e'(u^*)(\nabla y^* \cdot \nabla p^*) - F^* p^*, u - u^*) &\geq 0 \text{ for all } u \in U, \\ p^* \lambda^* = 0, \quad \mu^* y^* &= 0 \text{ a.e. in } \Omega, \\ (e(u^*)\nabla y^*, \nabla v) + (\lambda^*, v) - (Fu^*, v) &= 0 \text{ for all } v \in H_o^1(\Omega), \\ \Theta(\lambda^*, y^*) &= 0 \text{ a.e. in } \Omega, \end{aligned}$$

where  $\lambda^*$  is the weak limit of  $\{\max_{c_n}\{\bar{\lambda} - c_n \tilde{y}_{c_n}^*, 0\}\}$  in  $L^2(\Omega)$ , with  $\bar{\lambda}$  suitably chosen.

For the proof of this result we refer to [2]. In contrast to the first order characterizations in [1] and the references in [2,3], the above first order necessary condition is immediately amenable for numerical realization. Observe also the mixed nature of this system, i.e. partial differential equations and a variational inequality resulting from a partial differential equation are mixed with (algebraic) pointwise conditions. This mixture requires further investigation for numerical realization.



### 3. Numerical issues

For the discretization of the functions involved in the first order system of Theorem 4, we utilize finite dimensional subspaces and -sets yielding

$$y_h(x) = \sum_{i=1}^{N_y} Y_i w_i(x), \quad u_h(x) = \sum_{i=1}^{N_u} U_i u_i(x), \quad \lambda_h(x) = \sum_{i=1}^{N_\lambda} \Lambda_i l_i(x), \dots$$

We define  $\underline{Y} := (Y_1, \dots, Y_{N_y})^t \in \mathbb{R}^{N_y}$  and analogously for  $\underline{U}$ ,  $\underline{P}$ ,  $\underline{\Lambda}$ , and  $\underline{M}$ .

Let us briefly address some of the problems related to the pointwise conditions. For this purpose we assume that  $y$  is discretized by linear finite elements. If  $\lambda$  is discretized by piecewise constant finite elements, i.e. functions constant over each triangle, then one introduces linear dependencies. In fact, if  $Y_\ell > 0$ , then for all six triangles building the support for  $w_\ell(x)$  the coefficients  $\Lambda_j^i$ ,  $1 \leq j \leq 6$ , must vanish in order to satisfy  $\Theta(\lambda_h, y_h) = 0$ . On the other hand, even in cases where the infinite dimensional optimal solution satisfies strict complementarity, i.e.  $y(x) = 0$  implies  $\lambda(x) > 0$  for almost all  $x \in \Omega$ , in the discrete case we observe that at least one  $\Lambda_j^i$  may vanish although  $Y_\ell = 0$ . Both cases introduce numerical difficulties in terms of linearly dependent rows in the discretized first order system. Frequently, this results in instabilities.

For this reason we use the same number  $N$  of linear finite elements for  $y$ ,  $p$ ,  $\lambda$ , and  $\mu$ . For  $u$  we use  $N_u$  linear finite elements. A more detailed analysis and corresponding choices for the discretization will be reported elsewhere. As a consequence of this choice the pointwise almost everywhere conditions are enforced at the nodal points of the discretization. This implies that  $\Theta(\lambda, y)$  becomes  $\Theta_N(\underline{\Lambda}, \underline{Y}) = 0$  with  $\Theta_N(\underline{\Lambda}, \underline{Y}) = (\phi(\Lambda_1, Y_1), \dots, \phi(\Lambda_N, Y_N))^t$ , and  $\phi$  the complementarity function  $\phi(a, b) = \sqrt{a^2 + b^2} - (a + b)$  which satisfies  $\phi^2 \in C^1(\mathbb{R}^2; \mathbb{R})$ . In order to circumvent the variational inequality in the first order system we apply an interior technique for  $\underline{U}$ . More details on this aspect can be found in [2].

Another important issue comes from the fact that we cannot reliably estimate  $u$  on the singular set  $S_o = \{x \in \Omega | \nabla y(x) = 0\}$ . If we assume that  $y_d$  is a reasonable estimate for  $y$ , then we obtain a good approximation for  $S_o$ . On  $S_o$  we fix  $u|_{S_o} = \bar{u}|_{S_o}$ , where  $\bar{u}|_{S_o}$  is assumed to be available.

Since the resulting discretized first order system is overdetermined, i.e.  $5N + N_u$  equations have to be satisfied by  $4N + N_u - |S_{o,h}|$  variables, a stabilized Gauß-Newton approach is applied. For stabilization we use the forcing function  $\rho(r) = r^3$  with the norm of the relative residual as its argument. Globalization is achieved by an Armijo line search procedure.

Figures 4 – 6 give the result of a test run with  $\alpha = 10^{-4}$  for the lubrication problem. The dark gray part of the unit square represents the observation set, i.e. the set where  $u_h$  is computed. The algorithm needs 35 iterations to

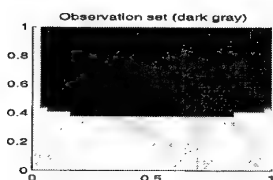


Fig. 4: Observation set.

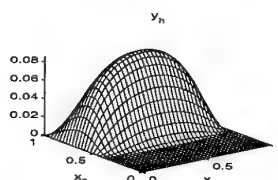


Fig. 5: Resulting pressure distribution  $y_h$ .

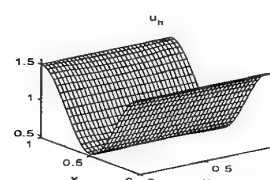


Fig. 6: Resulting height function  $u_h$ .

reduce the Gauß-Newton objective to the order of  $10^{-8}$ . The relative error in the max-norm for  $y_h$  is 0.142%. The corresponding error for  $u_h$  is 0.074%.

### 4. References

- 1 V. BARBU: *Optimal Control of Variational Inequalities*, Research Notes in Mathematics 100, Pitman.
- 2 M. HINTERMÜLLER: Inverse coefficient problems for variational inequalities: optimality conditions and numerical realization, SFB-report 174, Department of Mathematics, University of Graz, 1999.
- 3 K. ITO, AND K. KUNISCH: Optimal control of elliptic variational inequalities, Appl. Math. Optim. to appear.

Addresses: DR. MICHAEL HINTERMÜLLER, PROF. DR. KARL KUNISCH, University of Graz, Department of Mathematics, Heinrichstr. 36, A-8010 Graz, Austria.

KRAUSE ROLF H.; WOHLMUTH, BARBARA I.

## Domain Decomposition methods on nonmatching grids and some applications to linear elasticity problems

*Domain decomposition techniques provide a powerful tool for the coupling of different discretization methods or nonmatching triangulations across subregion boundaries. Here, we consider mortar finite elements methods for linear elasticity and diffusion problems. These domain decomposition techniques provide a more flexible approach than standard conforming formulations. The mortar solution is weakly continuous at subregion boundaries, and its jump is orthogonal to a suitable Lagrange multiplier space. Our approach is based on dual bases for the Lagrange multiplier space. It has the advantage of locally supported basis functions for the constrained space. This is not true for the standard mortar method [2]. The biorthogonality relation guarantees that the Lagrange multiplier can be locally eliminated, and that we obtain a symmetric positive semidefinite system on the unconstrained product space. This system will be solved by multigrid techniques. Numerical results illustrate the performance of the multigrid method in 2D and 3D.*

### 1. Introduction

The central idea of domain decomposition techniques is to decompose a global problem into subproblems of smaller complexity, and to "glue" the subproblems together in a suitable way. This is especially helpful for problems given on complicated geometries or problems with jumps in the material coefficients. Numerical examples for these situations will be given in the last section. As model problem let us consider the following linear elasticity problem with homogeneous Dirichlet boundary conditions

$$\sum_{j=1}^d \frac{\partial}{\partial x_j} \left( \sum_{l,m=1}^d E_{ijlm} \frac{\partial u_l}{\partial x_m} \right) = f_i \quad \text{in } \Omega.$$

Here,  $\Omega$  is a bounded, polygonal domain in  $\mathbb{R}^d$ ,  $d = 2, 3$  and Hooke's tensor  $E$  is assumed to be sufficiently smooth and uniformly positive definite. The components of  $f \in (L^2(\Omega))^d$  are denoted by  $f_i$ ,  $1 \leq i \leq d$ .

We use a geometrically conforming decomposition of  $\Omega$  into  $K$  non-overlapping polyhedral subdomains  $\Omega_k$ , i.e.,  $\bar{\Omega} = \bigcup_{k=1}^K \bar{\Omega}_k$ . Each subdomain is associated with an independent triangulations. Let us remark that the triangulations do not have to match at the common interface between two adjacent subdomains. The interfaces are denoted by  $\gamma_m$ ,  $1 \leq m \leq M$ , and inherit their triangulation from one of the adjacent subdomains. This side is called non-mortar side and the opposite one mortar side. The choice is arbitrary but fixed. We use standard piecewise linear in 2D and piecewise trilinear in 3D conforming finite elements on the subdomains and denote the product space by  $X_h$ . Then, the constrained mortar space  $V_h$  is defined by:  $V_h := \{v \in X_h \mid b(v, \mu) = 0, \mu \in M_h\}$ , where the bilinear form  $b(\cdot, \cdot)$  is given as a duality pairing on the interfaces  $b(v, \mu) := \sum_{m=1}^M \langle [v], \mu \rangle_{\gamma_m}$ ,  $v \in X_h$ ,  $\mu \in M_h$ , and  $[v]$  is the jump of  $v$  on  $\gamma_m$ . The constrained space  $V_h$  consist of all functions in  $X_h$  which have a vanishing jump at the interface with respect to the Lagrange multiplier space  $M_h$ . Of crucial importance is the suitable choice of  $M_h := \prod_{m=1}^M (M_h(\gamma_m))^d$  in the definition of the constrained space  $V_h$ . The space  $M_h(\gamma_m)$  is defined by  $\nu_m$  locally supported linear independent functions  $\psi_l^m$ . Here,  $\nu_m := \dim W_{0,h}(\gamma_m)$ , and  $W_h(\gamma_m)$  is the finite element trace space on  $\gamma_m$  and  $W_{0,h}(\gamma_m) := W_h(\gamma_m) \cap H_0^1(\gamma_m)$ . Now, the nonconforming mortar solution is obtained by: Find  $u_h \in V_h$  such that

$$a(u_h, v) = (f, v)_0, \quad v \in V_h. \quad (1)$$

Here, the bilinear form  $a(\cdot, \cdot)$  is given as  $a(u, v) := \sum_{k=1}^K \sum_{i,j,l,m=1}^d \int_{\Omega_k} E_{ijlm} \frac{\partial u_l}{\partial x_m} \frac{\partial v_i}{\partial x_j} dx$ ,  $u, v \in X_h$ . If  $P_0(\gamma_m) \subset M_h(\gamma_m)$ , the bilinear form  $a(\cdot, \cdot)$  is uniformly elliptic on  $V_h \times V_h$ . Under some assumption on  $M_h(\gamma_m)$ , optimal a priori bounds for the discretization error for the mortar finite element solution are obtained in the energy norm and the  $L^2$ -norm, we refer to [2,5,7].

**Dual bases for the Lagrange multiplier space.** Here, we consider two different types of dual bases in 2D and 3D. The first one is spanned by piecewise linear functions and the second one by piecewise constants. Figure 1 illustrates the shape of the dual basis functions. In the left part, the 2D case is depicted whereas in the right part, the isolines of a dual basis function at the two dimensional interface in 3D are given. In 2D, the support of our dual

basis functions is the union of exactly two adjacent edges, and in 3D it is in the interior of  $\gamma_m$  the union of four faces sharing one vertex. We remark that the definition of the basis functions associated with the vertex  $x_k^m$  has to be modified if  $x_k^m$  is close to the boundary of  $\gamma_m$ , for details we refer the reader to [2,7]. In both cases, it is easy to

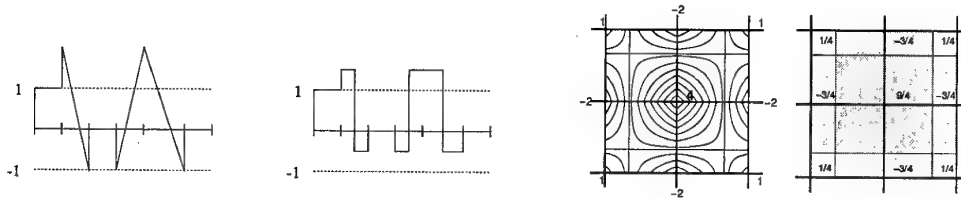


Figure 1: Piecewise constant and piecewise linear dual basis functions in 2D (left) and 3D (right)

see that the biorthogonality relation  $\int_{\gamma_m} \phi_l^m \psi_k^m d\sigma = \delta_{lk} \int_{\gamma_m} \phi_l^m d\sigma$  holds, where  $\phi_l^m$  and  $\psi_k^m$  denote the nodal basis functions of  $W_{0,h}(\gamma_m)$  and  $M_h(\gamma_m)$ , respectively. We refer to [7], for an analysis of the discretization error and some numerical results illustrating the influence of different Lagrange multiplier spaces. Of crucial importance is the so called mortar projection which is  $H_{00}^{1/2}$ -stable. Optimal a priori estimates in the energy norm and the  $L^2$ -norm can be obtained.

## 2. Multigrid method on the product space

Let us assume that we have a nested sequence of global triangulations and let us denote the associated unconstrained product spaces by  $X_l$ ,  $0 \leq l \leq L$ . The meshsize is given by  $h_l = 2h_{l+1}$ . In contrast to the constrained spaces  $V_l$ , the product spaces  $X_l$  are nested. We denote the standard prolongation operator by  $I_{l-1}^l : X_{l-1} \rightarrow X_l$  and the restriction by  $I_l^{l-1} : X_l \rightarrow X_{l-1}$ . Here, we use a symmetric positive semidefinite variational problem on the unconstrained product space  $X_l$  which is given in terms of a projection  $P_l$ . We start with the definition of the locally defined projection operator  $P_l : X_l \rightarrow X_l$

$$(P_l v)_i := \sum_{m=1}^M \sum_{l=1}^{\nu_m} \alpha_{m,l}^i \frac{b(v_i, \psi_l^m)}{\int_{\gamma_m} \phi_l^m d\sigma} \phi_l^m, \quad 1 \leq i \leq d,$$

where the nodal basis functions  $\phi_l^m$  of  $W_{0,l}(\gamma_m)$  are extended in a trivial way on the non-mortar side of  $\gamma_m$ . Then, it is easy to see that the kernel of  $P_l$  is exactly the constrained space  $V_l$ . Let  $A_l$ ,  $B_l$ , and  $C_l$  be the matrices associated with the bilinear forms  $a(\cdot, \cdot)$  on  $X_l \times X_l$ ,  $b(\cdot, \cdot)$  on  $X_l \times M_l$  and the projection  $P_l$ , respectively, and  $f_l$  the vector associated with the right hand side.

**Lemma 1.** *The following system is symmetric and positive definite. Its solution  $u_l \in X_l$  satisfies  $u_l \in V_l$  and (1)*

$$\hat{A}_l v_l := ((\text{Id} - C_l^T) A_l (\text{Id} - C_l) + C_l^T A_l C_l) u_l = (\text{Id} - C_l^T) f_l. \quad (2)$$

Furthermore, the solution  $u_l$  can be obtained by  $u_l = (\text{Id} - C_l) v_l$  from any solution  $v_l \in X_l$  of

$$\bar{A}_l u_l := (\text{Id} - C_l^T) A_l (\text{Id} - C_l) v_l = (\text{Id} - C_l^T) f_l. \quad (3)$$

**Proof.** It is easy to see that  $\hat{A}_l$  and  $\bar{A}_l$  are symmetric and positive semidefinite. Furthermore,  $A_l$  is positive definite on  $V_l$  and  $P_l X_l$ . Then, the triangle inequality yields that  $\hat{A}_l$  is positive definite. Now, let  $u_l \in V_l$  be the solution of (1), i.e.,  $(\text{Id} - C_l^T) A_l u_l = (\text{Id} - C_l^T) f_l$ , then by definition of  $C_l$  we find  $C_l u_l = 0$  and  $u_l$  solves (2). Observing that  $C_l$  is a projection, the second assertion follows immediately.

In the following, we call  $v_l \in X_l$  a solution of (3) only if it satisfies (3) and if  $v_l \in V_l$ . To obtain level independent convergence rates for our multigrid method, suitable approximation and smoothing properties have to be established. In a first step, we consider level dependent grid transfer operators  $(I_{\text{mod}})_l^{l-1}$  and  $(I_{\text{mod}})_{l-1}^l$  defined by

$$(I_{\text{mod}})_l^{l-1} := (\text{Id} - C_{l-1}^T) I_l^{l-1}, \quad (I_{\text{mod}})_{l-1}^l := (\text{Id} - C_l) I_{l-1}^l.$$

It is easy to see that these transfer operators guarantee  $C_{l-1}^T (I_{\text{mod}})_l^{l-1} w_l = 0$ ,  $w_l \in X_l$ , and  $C_l (I_{\text{mod}})_{l-1}^l w_{l-1} = 0$ ,  $w_{l-1} \in X_{l-1}$ . Then, an appropriate approximation property can be found in [9]. It is based on the assumption that the iterate after the smoothing steps is in the constrained space  $V_l$ . Starting with an arbitrary smoother for  $\hat{A}_l$ , we construct a modified one satisfying this condition. Let  $G_l^{-1}$  be a smoother for  $\hat{A}_l$ , e.g., a damped Jacobi method. Then, we define our modified smoother by  $\tilde{G}_l^{-1} := (\text{Id} - C_l) G_l^{-1} (\text{Id} - C_l^T)$ , and denote the iterates by  $y_l^i$  and  $\tilde{y}_l^i$ , respectively. The following lemma shows the relation between the two different iterates.

Lemma 2. Under the assumption  $\tilde{y}_l^0 = (\text{Id} - C_l)y_l^0$ , the iterates  $\tilde{y}_l^i$  can be obtained from  $y_l^i$  by a local post-processing step

$$\tilde{y}_l^i = (\text{Id} - C_l)y_l^i, \quad i \geq 1.$$

Furthermore, the smoothing and stability properties of  $\tilde{G}_l^{-1}$  are inherited from  $G_l^{-1}$ , i.e.,

$$\|\tilde{A}_l \tilde{e}_l^i\| = \|\tilde{A}_l e_l^i\|, \quad \|\tilde{e}_l^i\| \leq C \|e_l^i\|,$$

where  $\tilde{e}_l^i$  and  $e_l^i$  are the corresponding iteration errors and the constant  $C < \infty$  does not depend on the level  $l$ .

Proof. Observing the special structure of the right hand side  $d_l$  and  $\tilde{A}_l$ , we obtain by induction

$$\tilde{y}_l^{i+1} = \tilde{y}_l^i + (\text{Id} - C_l)G_l^{-1}(\text{Id} - C_l^T)(d_l - \tilde{A}_l \tilde{y}_l^i) = (\text{Id} - C_l)(y_l^i + G_l^{-1}(d_l - \tilde{A}_l y_l^i)) = (\text{Id} - C_l)y_l^{i+1}.$$

The second assertion follows from the observation that  $C_l$  is a scaled mass matrix, the norm of which is bounded independently of  $l$ .

Our multigrid method for the solution of (1) will be defined in terms of the equation (3), the given modified transfer operators, the smoother  $\tilde{G}_l^{-1}$ , the implementation of which is realized in terms of  $G_l^{-1}$  and one local post-processing steps at the end of the smoothing iterations. Then, we obtain level independent convergence rates for the  $\mathcal{W}$ -cycle provided that the number of smoothing steps is large enough.

### 3. Numerical results

Here, we consider some numerical results illustrating the performance of our multigrid method in 2D and 3D. Our multigrid method has been implemented for scalar problems and systems of equations for 2D and 3D in the finite element toolbox UG, see [1]. We apply nested iteration and use a tolerance of  $5 \cdot 10^{-8}$  for the norm of the residuum as stopping criterion for the iteration. Our first example is a 2D plane strain example with discontinuous coefficients, discretized by linear finite elements on triangles. The computational domain is depicted in the left picture in Figure 2, and consists of a nut and a wrench. Dirichlet boundary conditions are applied at the handle of the wrench, i.e.,  $u_1(x, y) = 1/3 \cdot \|m - (x, y)^T\| \cdot \sin(\alpha)$ ,  $u_2(x, y) = 1/3 \cdot \|m - (x, y)^T\| \cdot (1 - \cos(\alpha))$ , and homogenous Dirichlet conditions at the interior boundary of the nut. Here,  $m$  denotes the midpoint of the nut and we set  $\alpha = \pi/30$ . The interface is located at the contact area between the nut and the wrench. We use a  $\mathcal{W}(3, 3)$ -cycle with a symmetric Gauß-Seidel smoother accelerated by a stabilized biconjugate gradient method (bicgstab). Table 1 shows the required number of iterations on each level and the number of unknowns. As it can be seen, the number of iterations is independent of the level. The distorted grid scaled by a factor of 10 is shown in the second picture from the left in Figure 2. An adaptive refinement strategy has been used, controlled by a residual based error estimator for mortar finite elements.

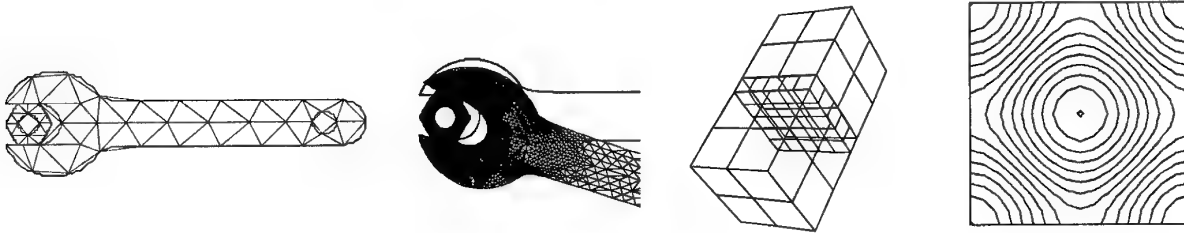


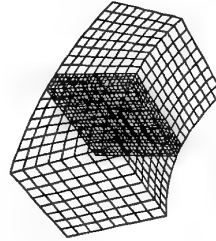
Figure 2: Initial and final triangulation in 2D (left) and initial triangulation and isolines in 3D (right)

As 3D example, we consider a "Sandwich"-like domain build up of two different materials. The domain  $\Omega$  is decomposed into three hexahedrons  $\bar{\Omega}_i := \{[0, 1]^2 \times [z_i, z_{i+1}]\}$  where  $z_1 := 0, z_2 := 1, z_3 := 1.2, z_4 := 2.2$ . In the right part of Figure 2, the nonmatching initial triangulation is shown. The non-mortar sides are defined on the middle hexahedron. We consider two different elliptic problems on this domain: a scalar model problem and a full linear elasticity problem, both with discontinuous coefficients. For both problems, we use the same initial triangulation, see Figure 2. We refer to the right picture in Figure 2 for the isolines of the solution at the interface in the scalar case.

Let us first consider the scalar problem  $-\text{div } a \nabla u = 1$ , on  $\Omega := (0, 1)^2 \times (0, 2.2)$  where the coefficient  $a$  is piecewise constant,  $a|_{\Omega_i} := 100, i = 1, 3$  and  $a|_{\Omega_2} := 1$ . Dirichlet boundary conditions are applied on the upper and lower part of the domain,  $u(x, y, z) = 1000 \sqrt{(x - 1/2)^2 + (y - 1/2)^2} \cdot (1.0 - y/3) \exp(-10(x^2 + y^2))$  if  $z = z_1$  or  $z = z_4$ , and homogeneous Neumann boundary conditions are taken on the remaining part of the boundary. In Figure 3, the

asymptotic convergence rates for the Jacobi and the Gauß–Seidel smoother are depicted. The numerical results show that the asymptotic convergence rates do not depend on the refinement level. Even for the  $\mathcal{V}(1, 1)$ -cycle, a constant asymptotic convergence rate is obtained. For the full linear elasticity example, we took as material parameters for

Level	# dof	# iter
0	108	1
1	232	3
2	904	4
3	1,622	4
4	2,350	4
5	3,478	5
6	5,380	5
7	8,272	5
8	12,844	5
9	20,130	5
10	30,878	5



Level	# dof	# iter
0	378	1
1	1,839	2
2	10,989	2
3	74,865	2
4	550,233	2

Table 1: Numerical results for the 2D example (left), the distorted grid (middle) and the 3D results (right)

the Lamé constants  $\mu|_{\Omega_i} = 8517$  and  $\lambda|_{\Omega_i} = 108280$  for  $i = 1, 3$  and  $\mu|_{\Omega_i} = 2008$  and  $\lambda|_{\Omega_i} = 3567$  for  $i = 2$ . Here, we apply an incomplete  $LU$ -decomposition as smoother and use the  $\mathcal{V}(3, 3)$ -cycle as preconditioner for the bcgstab-method. Dirichlet conditions are applied on the top and bottom of the "Sandwich", Neumann boundary conditions on the remaining part of the boundary. The right table in Table 1 shows the performance of our method in 3D. Although the number of unknowns increases by a factor of 10 in every refinement step, the number of iterations to achieve the required tolerance is constant. We remark, that uniform refinement has been used for this example. The displacement of the solution scaled by a factor of 10 is shown in the middle of Table 1.

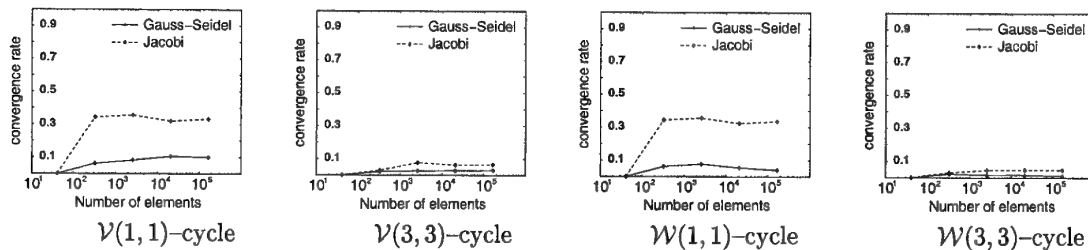


Figure 3: Asymptotic conv. rates for Jacobi and symmetric Gauß–Seidel smoother (3D scalar example)

#### 4. References

- 1 BASTIAN, P.; BIRKEN, K.; JOHANNSEN, K.; LANG, S.; NEUSS, N.; RENTZ-REICHERT, H.; AND WIENERS, C.: UG – a flexible software toolbox for solving partial differential equations. *Computing and Visualization in Science* 1 (1997), 27–40.
- 2 BERNARDI, C.; MADAY, Y.; AND PATERA, A.: Domain decomposition by the mortar element method, In H. Kaper et al., editors *Asymptotic and numerical methods for partial differential equations with critical parameters*. pages 269–286, 1993, Reidel, Dordrecht.
- 3 GOPALAKRISHNAN, G; AND PASCIAK, J.: Multigrid for the mortar finite element method. *to appear in SIAM J. Numer. Anal.*, 2000.
- 4 HACKBUSCH, W.: *Multi-Grid Methods and Applications*. Springer, 1985.
- 5 WOHLMUTH, B.: A mortar finite element method using dual spaces for the Lagrange multiplier. *to appear in SIAM J. Numer. Anal.*, 2000.
- 6 WOHLMUTH, B.; AND KRAUSE, R.: Multigrid methods based on the unconstrained product space arising from mortar finite element discretizations. *Preprint FUB A-18*, 1999.
- 7 WOHLMUTH, B.: Discretization Methods and Iterative Solvers Based on Domain Decomposition. *Habilitationsschrift*, Universität Augsburg, 1999.

Addresses: ROLF H. KRAUSE, Inst. für Mathematik, FU Berlin, 14 195 Berlin, Germany,  
email: krause@math.fu-berlin.de

BARBARA I. WOHLMUTH, Math. Institut, Universität Augsburg, 86 159 Augsburg, Germany,  
email: wohlmuth@math.uni-augsburg.de



ZUMBUSCH, G.

## On the quality of space-filling curve induced partitions

*The solution of partial differential equations on a parallel computer is usually done by a domain decomposition approach. The mesh is split into several partitions mapped onto the processors. However, partitioning of unstructured meshes and adaptive refined meshes in general is an NP-hard problem and heuristics are used. In this paper space-filling curve based partition methods are analysed and bounds for the quality of the partitions are given. Furthermore estimates for parallel numerical algorithms such as multigrid and wavelet methods on these partitions are derived.*

### 1. The partition problem

Finite-Element, Finite-Volume and Finite-Difference methods for the solution of partial differential equations are based on meshes. The solution is represented by degrees of freedoms attached to certain locations on the mesh. Numerical algorithms operate on these degrees of freedom during steps like the assembly of a linear equation system or the solution of an equation system. A natural way of porting algorithms to a parallel computer is the data distribution approach. The mesh with attached degrees of freedom is decomposed into several partitions and mapped to the processors of the parallel computer. Accordingly also the operations on the data are partitioned. Goals of a partitioning scheme are load-balancing and little communication between the processors. Sometimes also singly-connected partitions are required. If the partitions are determined during run-time, furthermore a fast partitioning scheme itself is sought. This is e.g. the case within adaptive mesh refinement of a PDE solver.

The partitioning problem in general is NP-hard [18]. There are many heuristics based on graph connectivity or geometric properties to address this problem [2,6,12,13,19]. In practice fast heuristics are known. However, there is not much known about general quality of these methods. In contrary there exist examples, where single heuristics give really bad results.

In this paper we analyse a specific geometry based heuristic based on space-filling curves. It is cheap and helps to simplify the implementation of parallel algorithms [9,15,16,17,20,23]. We are interested in bounds for the quality of the partitions. This will lead us to general estimates on the parallel performance of advanced numerical algorithms on these partitions.

### 2. Space-filling curves

First we have to define curves. The term curve shall denote the image of a continuous mapping of the unit interval to the  $\mathbb{R}^d$ . Mathematically, a curve is space-filling if and only if the image of the mapping does have a classical positive  $d$ -dimensional measure. The curve fills up a whole domain. For reasons of simplicity we restrict our attention to simple domains. We are interested in a mapping

$$f : [0, 1] \rightarrow I \mapsto \Omega \subset \mathbb{R}^d, \quad f \text{ continuous and surjective} \quad (1)$$

There are classical curves like the Hilbert-, the Peano- and the Lebesgue-curve, see [21]. However, we will also construct special space-filling curves on an unstructured mesh.

The space-filling curve can also be used for the inverse mapping  $f$  from a domain  $\Omega \subset \mathbb{R}^d$  to the unit interval  $I$ . This means that we can map geometric entities in  $\mathbb{R}^d$  to the one dimensional interval such as elements or nodes. Entities, which are neighbours on the interval, are also neighbours in the volume  $\mathbb{R}^d$ . Unfortunately the reverse cannot be true and neighbours in the volume may be separated through the mapping.

However, we can solve the resulting one-dimensional partition problem: We cut the interval  $I$  into disjoint sub-intervals  $I_j$  of equal workload with  $\bigcup_j I_j = I$ . This gives perfect load-balance and small separators between the partitions. The partition  $f(I_j)$  of the domain  $\Omega$  induced by the space-filling curve with  $\bigcup_j f(I_j) \supset \Omega$  also gives perfect load-balance. However, the separators  $\partial f(I_j) \setminus \partial \Omega$  are larger than the optimal separators in general as we will see.

### 3. Quality of a partition

We use a basic performance model for a distributed memory computer. The execution time of a program consists of computing time, which is proportional to the number of operations on a processor, and of communication time. Communication between the processors is done with message passing through some network and requires time linear in the size of data  $t = t_{\text{startup}} + n * t_{\text{bandwidth}}$ .

We consider  $O(n)$  algorithms linear in the size of data  $n$ , e.g. FEM matrix assembly for  $n$  finite elements, sparse matrix multiply or components of a multigrid algorithm such as a grid transfer or smoother, see [2,10]. The parallel computing time is  $C_1 \cdot n/p$  for a partition of  $n$  data onto  $p$  processors. We call  $v := n/p$  the *volume*. The runtime depends on the communication time. The data to be transferred is proportional to the separator or surface  $s_j$  of the partition  $s_j := \partial f(I_j) \setminus \partial \Omega$ .

$$t = C_1 \frac{n}{p} + C_2(t_{\text{startup}} + s * t_{\text{bandwidth}}) \quad (2)$$

This model suggest that we have to minimise the surface to volume ratio  $s/v$  of the partition for a high parallel efficiency of

$$\text{efficiency} = 1 / \left( 1 + \frac{C_2}{C_1} \left( \frac{1}{v} t_{\text{startup}} + \frac{s}{v} * t_{\text{bandwidth}} \right) \right). \quad (3)$$

While the lowest continuous surface to volume ratio is obtained for the sphere by  $s = \sqrt[2]{2d^{d-1} \frac{\pi^{d/2}}{\Gamma(d/2)}} v^{(d-1)/d}$ , we usually deal with partitions aligned with the mesh. Hence the cube with  $s = 2dv^{(d-1)/d}$  is of interest. In general we regard estimates of type

$$s \leq C_{\text{part}} \cdot v^{(d-1)/d} \quad (4)$$

with low constants  $C_{\text{part}}$  as optimal.

### 4. Estimates for space-filling curves

The estimate for the locality of a discrete space-filling curve  $F$  we will use with  $F : [1, \dots, k^d] \mapsto [1, \dots, k]^d$  is of type

$$\|F(x) - F(y)\|_2 \leq C \sqrt[d]{|x - y|}. \quad (5)$$

Gotsman and Lindenbaum [8] give an upper bound  $C = (d+3)^{d/2} 2^d$  for the Hilbert curve and tighter bounds for  $C = 6\frac{2}{3}$  for  $d = 2$  and  $C = 23$  for  $d = 3$ , which has been improved by [1]. Analogous estimates have been derived for the Hilbert curve [22] and the Peano curve [7]. It turns out that a similar curve, called H-index gives even better constants, see [5,14].

**Lemma 1.** *Given a connected discrete space-filling curve  $F$  on a domain  $[1, \dots, k]^d$  and a partition  $F([j, \dots, j+v-1])$  of  $v$  nodes, the surface  $s$  of the partition is bounded by eqn. 4. The constant  $C_{\text{part}}$  depends on the curve.*

**Proof.** is based on eqn. 5 and the connectedness of the partition. It is sufficient to consider  $s$  of the bounding box.

This lemma does not hold for curves of Lebesgue also called *bit-interleaving* [3], because the discrete partitions tend to be disconnected. However, we generalise the situation to unstructured and adaptively refined meshes by the following construction: We create an enumeration of a mesh by some heuristic in order to obtain a 'local' discrete space-filling curve. Then we do mesh refinement by some geometric refinement rules, see [2,4]. Each coarse element  $E_j$  is substituted by several smaller  $E_{j,k}$  elements. The enumeration is changed such that it cycles through these new elements  $E_{j,k}$  right after the elements  $E_{j-1}$  or  $E_{j-1,k}$ . This leads in the limit to a continuous space-filling curve, see [11,15,20]. Alternatively a standard, continuous space-filling curve can be super-imposed onto the grid, see [9,17].

**Corollary 2.** *Estimate 4 also holds for a space-filling curve partitioning of a (quasi-) uniform mesh by superposition of  $f$  or mesh dependent construction of  $f$ .*

Estimate 3 combined with corollary 2 gives a speedup for large problems of

$$\text{efficiency} = 1 / \left( 1 + \frac{C_2 C_{\text{part}} t_{\text{bandwidth}}}{C_1} \cdot \frac{p}{n^{1/d}} \right). \quad (6)$$

This implies optimal parallel efficiency for very large problems,  $n \rightarrow \infty$ . Estimate 6 holds for a code for the solution of partial differential equations in the steps of setting up an equation system, a single matrix multiply, a fixed number of Krylov iterations. Furthermore, using the same space-filling curve on all grid levels, this also holds for an additive multigrid implementation and for standard multigrid if we neglect terms  $\log n \cdot t_{\text{startup}}$  proportional to the number of grid levels. For the scalability of a global PDE solver an  $O(n)$  multigrid solver is essential. Solvers with higher than linear complexity may scale in  $p$  like eqn. 6 but scale completely different in  $n$ .

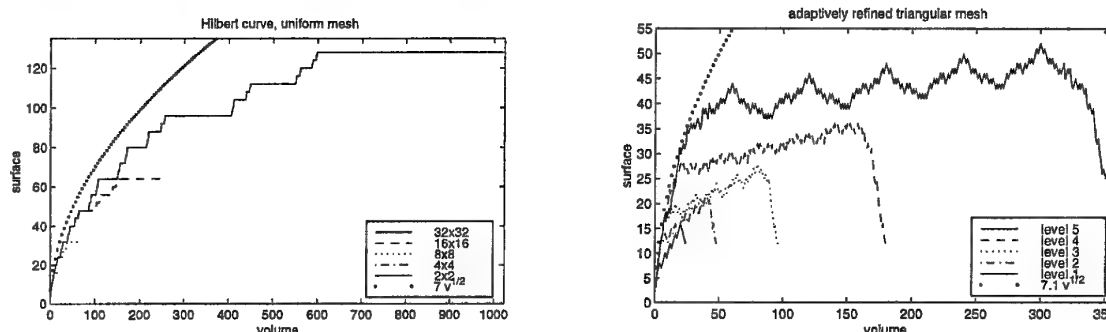


Figure 1: Locality of partitions defined by a space-filling curve. Hilbert curve on an uniform mesh (left) and an unstructured mesh with adaptive mesh refinement (right).

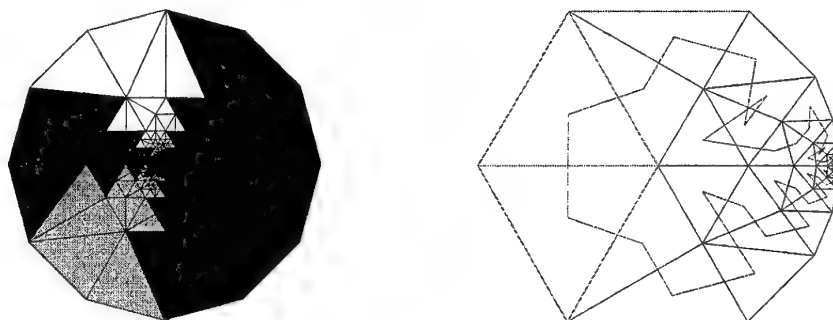


Figure 2: Adaptive mesh refinement. Partitions defined by a space-filling curve (left) and a counter example for a non-local partition (right).

## 5. Numerical experiments

The proof of lemma 1 only gives a crude estimate on the constant in eqn. 4. Hence we look at two examples for two-dimensional partitions. In figure 1 the maximum surfaces  $s$  to different volumes  $v$  are given. We consider a uniform square  $[0, 2^k]^2$  (counting the complete boundary) and a triangulation (counting the interior boundary only). The triangulation starts with a hexagon and angles of  $\pi/3$  and is refined adaptively. The triangulation is shown in figure 2 left. The different graphs in figure 1 show the ratios for different grid levels. The surface of small partitions comes close to the expected  $\sqrt{n}$  behaviour while larger partitions have a limited boundary.  $\partial\Omega$  is a natural limit.

Lemma 1 did not deal with adaptive mesh refinement. Although, moderate refinement seems to give similar estimates. However, very strong refinement with an arithmetic progression of nodes during refinement shows a different picture. In this example, figure 2 right,  $s$  is proportional to  $v$ . This behaviour limits the usefulness of the partition method. This 'counter' example is related to examples where other heuristics like spectral bisection [19] also fail to perform well.

## 6. Sparse grids

Space-filling curve partitions can also be used for the parallelisation of adaptive sparse grid implementations, see [25]. A certain choice of tensor products of (pre-) wavelet basis functions can give approximations with a low number of degrees freedom of the order  $v = O(n \log^{d-1} n)$  for a spatial resolution of  $1/n$ , see [24]. The discretization of PDEs on such sparse grids links geometrical nodes on different scales and at different distances. The surface of a rectangular shaped partition is of order  $s = O(n \log^{d-2} n)$  which is rather large compared to  $v$ . Experimentally space-filling curves and other graph partitions heuristics give partition surfaces of similar size.

$$s \leq C_{\text{part}} \frac{v}{\log n} \quad \text{with } v = O(n \log^{d-1} n) \quad (7)$$

$$\text{efficiency} = 1 / (1 + C \frac{p}{\log n}) \quad (8)$$

We obtain scalability of wavelet algorithms on sparse grids. However, the parallel efficiency grows far slower in the problem size as for standard discretizations which scale excellently, compare eqns. 6 and 8.

## 7. References

- 1 ALBER, J., NIEDERMEIER, R.: Hilbert indexings in arbitrary dimensions. Theory of computing systems, to appear.
- 2 BASTIAN, P.: Parallele Adaptive Mehrgitterverfahren. Teubner (1996).
- 3 BEICHL, I., SULLIVAN, F.: Interleave in Peace, or Interleave in Pieces. IEEE Computational Science Engineering 2, (1998), 92–96.
- 4 BEY, J.: Tetrahedral grid refinement. Computing 55, (1995), 355–378.
- 5 CHOCHIA, G., COLE, M.: Recursive 3D mesh indexing with improved locality. in: Proc. HPCN '97, LNCS 1225, Springer (1997).
- 6 DUBINSKI, J.: A parallel tree code. New Astronomy 1, (1996), 133–147.
- 7 GARSIA, A. M.: Combinatorial inequalities and smoothness of functions. Bull. Amer. Math. Soc. 82 (1976), 157–170.
- 8 GOTSMAN, C., LINDENBAUM, M.: On the metric properties of discrete space-filling curves. IEEE Trans. on Image Processing 5 (1996), 794–797.
- 9 GRIEBEL, M., ZUMBUSCH, G.: Hash-Storage Techniques for Adaptive Multilevel Solvers and their Domain Decomposition Parallelization. in: Proc. Domain Decomposition Methods 10. Mandel, J., Farhat, C., Cai, X.-C., Contemporary Mathematics 218, AMS (1998), 279–286.
- 10 GRIEBEL, M., ZUMBUSCH, G.: Parallel adaptive subspace correction schemes with applications to elasticity. Computer methods in applied mechanics and engineering. 184 (2000), 303–332.
- 11 HEBER, G., GAO, G. R., BISWAS, R.: Self-Avoiding Walks over Adaptive Unstructured Grids. Concurrency - Practice and Experience. 12 (2000), 85–109.
- 12 KARYPIS, G., KUMAR, V.: Multilevel k-way Graph Partitioning for Irregular Graphs. J. Parallel and Distributed Comp. 48 (1998), 96–129.
- 13 DE KEYSER, J., ROOSE, D.: Partitioning and Mapping Adaptive Multigrid Hierarchies on Distributed Memory Computers. Report TW 166, Univ. Leuven, Dept. Computer Science (1992).
- 14 NIEDERMEIER, R., REINHARDT, K., SANDER, P.: Towards optimal locality in mesh-indexings. in: Proc. 11th Int. Symp. Fundamentals of Computation Theory, LNCS 1279 (1997), 364–375.
- 15 ODEN, J. T., PATRA, A., FENG, Y.: Domain Decomposition for Adaptive hp Finite Element Methods. in: Proc. Domain Decomposition 7, Contemporary Mathematics 180, AMS (1994), 295–301.
- 16 PARASHAR, M., BROWNE, J. C.: On partitioning dynamic adaptive grid hierarchies. in: Proc. 29th Annual Hawai Int. Conf. System Sciences (1996).
- 17 PILKINGTON, J. R., BADEN, S. B.: Partitioning with Spacefilling Curves. Report CS94-349, UCSD, Dept. Computer Science (1994).
- 18 POTHEN, A.: Graph Partitioning Algorithms with Applications to Scientific Computing. in: Parallel Numerical Algorithms, Keyes, D. E., Sameh, A., Venkatakrishnan, V. (eds.), Kluwer (1997), 323–368.
- 19 POTHEN, A., SIMON, H., LIOU, K.-P.: Partitioning Sparse Matrices with Eigenvectors of Graphs, SIAM J. Matrix Analysis and Applications 11 (1990), 430–452.
- 20 ROBERTS, S., KALYANASUNDARAM, S., CARDEW-HALL, M., CLARKE, W.: A Key Based Parallel Adaptive Refinement Technique for Finite Element Methods, in: Proc. CTAC97, Noye, B. J., Teubner, M. D., Gill, A. W. (eds.), World Scientific (1998).
- 21 SAGAN, H.: Space-Filling Curves, Springer (1994).
- 22 STOUT, Q.F.: Topological matching. Proc. 15th ACM Symp. on Theory of Computing (1983), 24–31.
- 23 WARREN, M., SALMON, J.: A portable parallel particle program. Comput. Phys. Comm. 87, 266–290.
- 24 ZENGER, C.: Sparse Grids. in: Proc. 6th GAMM Seminar, Hackbusch, W. (eds.), Vieweg (1991).
- 25 ZUMBUSCH, G.: A Sparse Grid PDE Solver. in: Advances in Software Tools for Scientific Computing, Langtangen, H. P., Bruaset, A. M., Quak, E. (eds.), LNCSE 10, Springer (2000), 133–177.

Addresses: DR. GERHARD ZUMBUSCH, Institut für Angewandte Mathematik, Universität Bonn, D-53115 Bonn Germany

DENZLER, JOCHEN

## Branch Classification for the Spectrum of the Orr–Sommerfeld Equation for Plane Couette Flow

*The spectrum of the Orr–Sommerfeld equation for Plane Couette Flow (solution continuum of the characteristic equation) undergoes only remarkably simple topological changes as the Reynolds number changes from 0 to  $\infty$ . The case of Reynolds number 0, which can be treated completely in full rigor, determines most of the structure for any finite Reynolds number, but not for infinite Reynolds number, the only complementary phenomenon being the so-called mode crossing.*

### 1. Introduction

The purpose of this paper is to give a detailed rigorous branch classification of the eigenvalue spectrum of the Orr–Sommerfeld equation

$$(\mathcal{R}^{-1}(\partial_y^2 - \alpha^2) - i\alpha y - \lambda)(\partial_y^2 - \alpha^2)w(y) = 0, \quad w(0) = 0 = w(1), \quad w'(0) = 0 = w'(1) \quad (1)$$

for plane Couette flow (pCf). This equation arises as the linearization of the Navier–Stokes equation for a plane viscous incompressible shear flow between two parallel plates moving with respect to each other, as is detailed out, e.g., in [2],[4]. It has been shown by Romanov [3] that, for any Reynolds number  $\mathcal{R}$  and any wave number  $\alpha$ , all eigenvalues satisfy  $\operatorname{Re} \lambda \leq -\mu < 0$ , and also that asymptotic stability of the solution under the full Navier–Stokes equations ensues from this in the given case.

Given such a result, one still has to beware that the basin of attraction of the asymptotically stable solution may be very small, and it is in fact generally believed to shrink indefinitely as  $\mathcal{R} \rightarrow \infty$  in the case of pCf. This explains superficially why instability is observed at large  $\mathcal{R}$  in spite of Romanov's stability result. A more detailed account of the phenomenon is lacking. Trefethen and others [5] are working out scenarios where the non-normality of the linearized operator accounts for a large transient growth of initial perturbations, which will therefore be driven out of the basin of attraction by this effect. Much of that work is numerical, e.g. [6], or treats a model problem rather than the original equation [1], owing to the inherent difficulty. Any deeper understanding of the dynamics in a neighbourhood of the shear flow clearly requires not merely the half plane estimates for the spectrum, but a closer understanding of both the eigenvalues and the eigenfunctions coming with them (which amounts to a similar information as is contained in the pseudospectrum). This work gives rigorous results for the eigenvalues. The eigenfunctions will require further study, which is on the way. We give rough sketches of proofs only. Details will appear elsewhere.

### 2. The characteristic equation

By means of the transformation

$$2t = (\alpha\mathcal{R})^{1/3}, \quad \lambda = -\mathcal{R}^{-1}(4it^3 + 4t^2\delta) - \mathcal{R}^{-1}(8\mathcal{R}^{-1}t^3)^2, \quad \varepsilon = \mathcal{R}^{-1} \quad (2)$$

the eigenvalue condition of (1) can be written as  $\Delta(\delta, t, \varepsilon) = 0$  with the well-known Gram determinant

$$\Delta(\delta, t, \varepsilon) := \begin{vmatrix} \int_{-1}^1 \operatorname{Ai}(\delta + itz) \sinh(4\varepsilon t^3 z) dz & \int_{-1}^1 \operatorname{Ai}(\delta + itz) \cosh(4\varepsilon t^3 z) dz \\ \int_{-1}^1 \operatorname{Bi}(\delta + itz) \sinh(4\varepsilon t^3 z) dz & \int_{-1}^1 \operatorname{Bi}(\delta + itz) \cosh(4\varepsilon t^3 z) dz \end{vmatrix}. \quad (3)$$

It is sufficient to study this equation for  $t > 0$  and complex  $\delta$ . Simplifications of (3) occur in various limiting cases. The case  $\mathcal{R} \rightarrow 0$ ,  $\varepsilon \rightarrow \infty$ , physically irrelevant as it may seem, leads to  $\frac{\operatorname{Bi}}{\operatorname{Ai}}(\delta + it) = \frac{\operatorname{Bi}}{\operatorname{Ai}}(\delta - it)$  and can be handled comparatively easily and elegantly. It turns out that it not only gives the branch structure for sufficiently small  $\mathcal{R}$ , but actually up to arbitrarily large  $\mathcal{R}$  with exception of only finitely many (depending on  $\mathcal{R}$ ) branches. The limiting case  $\mathcal{R} \rightarrow 0$  is therefore a good starting point for the branch classification by homotopy arguments, and we shall give details below.

### 3. Series expansion

There is a remarkable series representation of (3) that seems to be completely new:

**Theorem 1.** Let  $x := 4\delta t^2$  and  $A := 64\varepsilon^2 t^6$ . Then we have the locally, in  $(\mathbb{C} \setminus \{0\}) \times \mathbb{R} \times [0, \infty[ \ni (\delta, t, \varepsilon)$ , uniformly absolutely convergent series expansion

$$-\pi \frac{\Delta(\delta, t, \varepsilon)}{16i\varepsilon t^4} = \sum_{j=0}^{\infty} \frac{\delta^{-3j}}{(A+x)^{4j+2}} \left[ P_j^{00}(x, A) + P_j^{cc}(x, A) \cos \sqrt{x} \cosh \sqrt{A} + \right. \\ \left. + P_j^{cs}(x, A) \cos \sqrt{x} \frac{\sinh \sqrt{A}}{\sqrt{A}} + P_j^{sc}(x, A) \frac{\sin \sqrt{x}}{\sqrt{x}} \cosh \sqrt{A} + P_j^{ss}(x, A) \frac{\sin \sqrt{x}}{\sqrt{x}} \frac{\sinh \sqrt{A}}{\sqrt{A}} \right] \quad (4)$$

with polynomials  $P_j^{**}$  consisting of monomials of which the joint degrees with respect to  $(x, A)$  are within the range given by the table:

deg of mon's in	$P_j^{00}$	$P_j^{cc}$	$P_j^{cs}$	$P_j^{sc}$	$P_j^{ss}$
... from	$4j$	$4j$	$4j+1$	$4j+1$	$4j+1$
... to	$5j$	$5j + \lfloor j/2 \rfloor$	$5j + \lfloor (j+1)/2 \rfloor$	$5j + \lfloor (j+1)/2 \rfloor$	$5j + \lfloor j/2 \rfloor + 1$
deg wrt $A \leq$	$2j$	$4j$	$4j+1$	$4j$	$4j+1$

The apparent pole at  $x = -A$  is compensated, in each order  $j$ , by a zero of order at least  $4j+2$  of the term in brackets.

The theorem is proved by means of a full Taylor expansion with respect to  $t$  of all terms under the integrals in (3). A lot of combinatorics is involved to recombine the terms into series (4). Explicit (though rather unwieldy) combinatorial formulas for the polynomials are available.

In the case  $t < O(|\delta|^{1/4})$ , the  $j = 0$  term of (4) in Theorem 1 is actually dominant and permits a full classification of branches in that domain:

**Theorem 2.** Given  $\varepsilon \geq 0$  and  $C$  sufficiently small, independent of  $\varepsilon$ , the zero set

$$\{(\delta, t) \in \mathbb{C} \times \mathbb{R}^+ \mid \varepsilon^{-1} \Delta(\delta, t, \varepsilon) = 0, t < C|\delta|^{1/4}\} \quad (5)$$

(with the obvious interpretation of the case  $\varepsilon = 0$  as a limit  $\varepsilon \rightarrow 0$ ) consists of real branches  $(\delta_n(t), t)$ ,  $n \geq 2$ , only. These branches are totally ordered:  $\delta_n(t) < \delta_{n+1}(t)$ , the limit  $\omega_n := \lim_{t \rightarrow 0} 2\sqrt{\delta_n(t)} t$  exists and is independent of  $\varepsilon \in [0, \infty[$ . It holds

$$(n-1)\pi < 2\sqrt{\delta_n(t)} t < \begin{cases} (n + \frac{1}{2})\pi & \text{for } n \text{ even} \\ n\pi & \text{for } n \text{ odd} \end{cases}, \quad \omega_n = n\pi \quad \text{for } n \text{ even} \\ \omega_n/2 = \tan \omega_n/2 \quad \text{for } n \text{ odd} \quad (6)$$

For  $\varepsilon = \infty$  (where in (5) the corresponding limiting equation  $\frac{\text{Bi}}{\text{Ai}}(\delta + it) = \frac{\text{Bi}}{\text{Ai}}(\delta - it)$  appears), it holds  $\lim_{t \rightarrow 0} 2\sqrt{\delta_n(t)} t = (n-1)\pi$ . For arbitrary  $C$ , the same still applies to those branches indexed by  $n \geq n_0(C)$ .

These real solutions are the basis of the branch classification. Based on either the implicit function theorem (IFT) or, where  $\partial_\delta \Delta = 0$ , on the Weierstraß preparation theorem (WPT), any solution to  $\Delta(\delta, t, \varepsilon) = 0$  can be followed for ever decreasing  $t > 0$ , until it finally enters the domain covered by Theorem 2: In points where  $\partial_\delta \Delta = 0$ , the IFT fails, but WPT shows that a branch point does actually occur, and one can follow the solution continuum for further decreasing  $t$  (nonuniquely, though, with a choice of branches). Therefore all solution branches are connected to the real branches discussed in Theorem 2 through (possibly several) bifurcations.

The proof of Theorem 2 is a Rouché counting argument, combined with an existence proof for real solutions by the intermediate value theorem, where the real solutions thus found already exhaust the total number of solutions determined by Rouché.

### 4. Small or finite Reynolds numbers

The behaviour outside the domain  $t < O(|\delta|^{1/4})$  is determined by the following theorem:

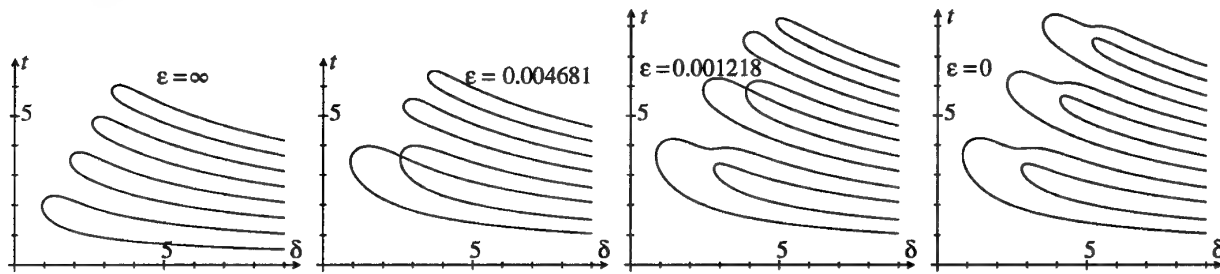
**Theorem 3.** For sufficiently large  $\varepsilon$  (small Reynolds number), the zero set  $\mathcal{Z}_\varepsilon^+ := \{(\delta, t) \mid t > 0, \Delta(\delta, t, \varepsilon) = 0\}$  is homeomorphic to  $\mathcal{Z}_\infty^+ := \{(\delta, t) \mid t > 0, \frac{\text{Bi}}{\text{Ai}}(\delta + it) = \frac{\text{Bi}}{\text{Ai}}(\delta - it)\}$ , namely it consists of

- (1) A family of real branches  $\tilde{R}_j$ ,  $j \geq 1$ , connecting the  $(\delta_{2j}(t), t)$  and  $(\delta_{2j+1}(t), t)$  of Theorem 2 into a single curve on which the coordinate  $t$  has exactly one critical point, a nondegenerate maximum. (Call this point  $r_j$ .)
- (2) A corresponding branch  $\tilde{C}_j$  of complex solutions  $(\delta_j(t), t)$  and  $(\bar{\delta}_j(t), t)$  branching off at  $r_j$  for increasing  $t$  and globally parametrized by  $t$ . As  $t \rightarrow \infty$ ,  $(\delta_j(t) - it)e^{-2\pi i/3}$  converges to  $a_{-j} < 0$ , the  $j^{\text{th}}$  zero of the Airy function  $\text{Ai}$  (independent of  $\varepsilon$ ).

For arbitrary  $\varepsilon > 0$  (not necessarily large), the same applies to those branches indexed by  $j \geq j_0(\varepsilon)$ .

Given this result, it follows that changes in the topology of  $Z_\varepsilon^+$  can only accumulate as  $\varepsilon \rightarrow 0$  ( $\mathcal{R} \rightarrow \infty$ ) with  $t \rightarrow \infty$  at the same time. A quantitative analysis shows that they actually are confined to a critical domain where  $|t - \sqrt{3}\delta| < O(t^{-1/2} \ln t)$  as  $t \rightarrow \infty$ . By means of a separate calculation for the case  $\varepsilon = 0$  and comparison to  $\varepsilon > 0$  one sees that bifurcations actually do occur for all branches (see Theorem 4 below).

The figure shows the first few real branches for different values of  $\varepsilon$ . The asymptotic behaviour as  $t \rightarrow 0$  is clearly visible, both for  $\varepsilon < \infty$  as stated in (6) and for  $\varepsilon = \infty$ , where the asymptotic behaviour is shifted. The change of topology which affects the first real branches, one at a time, will be discussed in the following section.



## 5. The inviscid limit

Here the following branch classification holds. (Remember that the  $\delta_n$  of Theorem 2 are numbered from  $n = 2$ .)

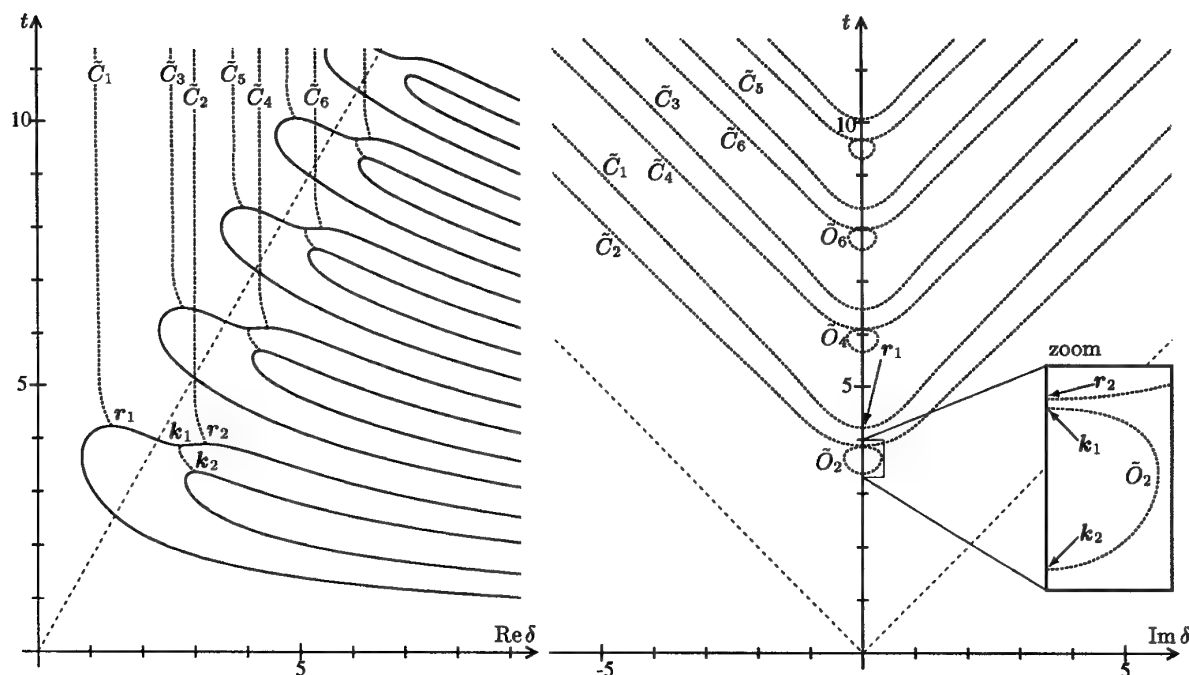
**Theorem 4.** The solution continuum  $Z_0^+$  consists of

- (1a) real branches  $\tilde{R}_{2j}$ ,  $j$  sufficiently large, connecting the  $(\delta_{4j-1}(t), t)$  and  $(\delta_{4j}(t), t)$  of Theorem 2 into a single curve on which the coordinate  $t$  has exactly one critical point, a nondegenerate maximum. (Call this point  $k_{2j}$ .)
- (1b) real branches  $\tilde{R}_{2j-1}$ ,  $j$  sufficiently large, connecting the  $(\delta_{4j-2}(t), t)$  and  $(\delta_{4j+1}(t), t)$  of Theorem 2 into a single curve on which the coordinate  $t$  has exactly three critical points, a nondegenerate minimum  $k_{2j-1}$  and two nondegenerate maxima  $r_{2j-1}$  and  $r_{2j}$ .
- (2) complex branches  $\tilde{C}_{2j-1}$  and  $\tilde{C}_{2j}$  bifurcating from  $\tilde{R}_{2j-1}$  and  $\tilde{R}_{2j}$  at  $r_{2j-1}$  and  $r_{2j}$  respectively. These consist of pairs of complex conjugate solutions parametrized globally by  $t$  as in Theorem 3. This time, however,  $(\delta_j(t) - it)e^{-2\pi i/3}$  converges to zeros of  $\int^\infty \text{Ai}(x) dx$  instead, as  $t \rightarrow \infty$ .
- (3) topological circles  $\tilde{O}_{2j}$  connecting  $k_{2j-1}$  with  $k_{2j}$ , and also consisting of pairs of complex conjugate solutions parametrized by  $t$ .

Numerical evidence suggests that “ $j$  large enough” means actually “ $j \geq 1$ ”, thus exhausting all branches. Without resort to numerics, all asymptotic statements hold for  $j \geq 1$ ; it is only the way how the ends are connected inside a compact set that is affected by the hypothesis “ $j$  large enough”.

The proof of this theorem relies on asymptotic expansions of Airy functions. It is very likely possible, by enhancing the qualitative asymptotic estimates entering in the proof with quantitative error bounds, to abandon the hypothesis “ $j$  large” in favor of “ $j \geq 1$ ”. We have however not carried this through. Numerical evidence also suggests that the mode crossings that  $\tilde{R}_{2j-1}$  and  $\tilde{R}_{2j}$  undergo as  $\varepsilon \rightarrow 0$  are the only bifurcations to occur, and that they occur one after another for  $j = 1, 2, 3, \dots$ , with no changes of topology being subsequently undone again. A rigorous proof of this at least for large  $j$  is currently under investigation.

The figure shows the first few real and complex branches for  $\varepsilon = 0$  ( $\mathcal{R} = \infty$ ), projected in either the  $(\operatorname{Re} \delta, t)$  or the  $(\operatorname{Im} \delta, t)$  plane. The complex branches are dotted. The dashed lines ( $t = \sqrt{3}\delta$  and  $t = |\operatorname{Im} \delta|$ ) help to display the asymptotic behaviour.



## 6. Outlook

Next, a description of the corresponding eigenfunctions would be desirable. The method used to study the spectrum has this perspective in mind. In certain limiting cases, the eigenfunctions can be understood more easily: in particular, under the limit  $t \rightarrow 0$ ,  $2\sqrt{\delta}t \rightarrow \omega_n$ , the  $\alpha y$  term in (1) can be neglected, and the limiting eigenfunctions can be chosen as real. It is hoped that the classification of branches permits continuity methods to be applied to the eigenfunctions.

## Acknowledgements

*This work is part of my thesis of Habilitation at the Technische Universität München. It is a pleasure to acknowledge the continuing support of Professor Scheurle who opened this new field to me.*

## 7. References

- 1 J. S. BAGGETT, T. A. DRISCOLL, AND L. N. TREFETHEN: A mostly linear model of transition to turbulence," *Physics of Fluids* 7 (1995), 833-838.
- 2 P.G. DRAZIN, W.H. REID: *Hydrodynamic Stability*. Cambridge Univ Pr, 1981
- 3 V.A. ROMANOV: Stability of Plane-Parallel Couette Flow. *Functional Analysis and Applications* 7(1973), 137-146 (translated from: *Funktsional'nyi analiz i ego prilozheniya* 7,2(1973), 62-73)
- 4 H.B. SQUIRE: On the Stability for Three-Dimensional Disturbances of Viscous Fluid Flow between Parallel Walls. *Proc. of the Royal Soc. of London, Ser. A* 142(1933), 621-628
- 5 LLOYD N. TREFETHEN, ANNE E. TREFETHEN, SATISH C. REDDY, TOBIN A. DRISCOLL: Hydrodynamic Stability Without Eigenvalues. *Science* 261, 30 July 1993, 578-584
- 6 ANNE E. TREFETHEN, LLOYD N. TREFETHEN, PETER J. SCHMID: Spectra and Pseudospectra for Pipe Poiseuille flow. *Comp. Meth. Appl. Mech. Engr.* 1926 (1999), 413-420.

*Address:* DR. JOCHEN DENZLER, Zentrum Mathematik (SCS), Technische Universität München, 80290 München, Germany. — As of August 2000: Department of Mathematics, University of Notre Dame, Notre Dame, IN 46556, USA



SERRE E., RASPO I., BONToux P. AND PEYRET R.

## Spectral Solution of the Navier-Stokes Equations for Rotating Flows

*Direct numerical simulation by spectral methods are developed and used to study the instabilities in the Ekman and Bödewadt layers at the transition to the time-dependent regimes. The physical phenomena are characteristic of rotating flows with walls. The geometrical cavities are elementary geometries that are relevant of turbine applications and that also refer to typical configurations studied in fundamental investigations and in experiments.*

*The three - dimensional Chebyshev - Fourier collocation method is based on a projection scheme to solve the coupling between the velocity and the pressure. The method devoted to annular domain is extended to fully cylindrical domain involving the axis of rotation. A special development has been required to deal with the singular behaviour of the coefficients when the radius tends to zero. The investigation was carried out in two types of generic configurations that are the forced throughflow in a rotating cavity and the confined flow driven by the differential rotation inside a rotor-stator cavity. Depending on the aspect ratio and on the Reynolds number, counter-rotating rolls can superimpose to the boundary layer flow near the disks under the form of annular and spiral structures. The characteristic parameters of the perturbations (wavelength, frequency, phase velocity, inclination of the spiral ) are shown to be relevant of the types I and II instabilities in rotating flows.*

## 1 Introduction

Instability patterns are simulated with an efficient spectral method in different rotating flows driven by the differential rotation of the walls or by a throughflow. Two enclosed rotor-stator annular and cylindrical cavities and a rotating cavity with a superposed radial outflow are considered. These situations are relevant of air cooling devices of gas turbine engines. A characteristic of these flows resides in the coexistence of adjacent and coupled flow regions that are greatly different in terms of the flow properties and of the length scales as it is the case for the Ekman or the Bödewadt layers and the geostrophic core region.

The characteristic parameters of the solutions are shown to be relevant with available theoretical and experimental results about type I and type II instabilities, that are two classes of generic instabilities developing above rotating disks [1]. The spatial structure of these instabilities consists of circular and spiral vortices on the boundary layers.

## 2 Geometrical and mathematical models

The geometrical models correspond to two disks of depth radius  $\Delta R = R_1 - R_0$ , where  $R_0$  and  $R_1$  are the internal and external radii. The geometrical domain can be open including radial inflow and outflow boundary conditions, or completely enclosed by one or two cylinders of height  $2h$ , the internal shaft and the external shroud. The open cavity rotates at the uniform angular velocity  $\Omega$  whereas the enclosed cavity is composed of a stationary part (stator) and of a rotating disk (rotor) at the angular velocity  $\Omega$ .

The motion is governed by the three-dimensional Navier-Stokes equations written in primitive variables for an incompressible fluid. The two geometrical parameters refer to curvature and the aspect ratio,  $R_m = (R_1 + R_0)/\Delta R$  and  $L = \Delta R/2h$ . The scales for the dimensionless variables of space, time and velocity are  $[h, \Omega^{-1}, \Omega R_1]$  respectively. The dimensionless spatial variables are denoted  $(\bar{r}, \bar{z})$  and have been normalized to the square form  $[-1, 1] \times [-1, 1]$ , a requirement for the use of Chebyshev polynomials. The normalized spatial variables are denoted  $(r, z)$  and  $r = (2h\bar{r} - a - b)/\Delta R$  and  $z = \bar{z}$ . The relevant physical parameters are the Reynolds number, defined as  $Re = \Omega (2h)^2 / \nu$  and when the flow is submitted to a radial outflow, the mass flow rate,  $Q$ , is made dimensionless as  $C_w = Q/\nu R_1$ . We note by radial outflow that the flow is forced from  $R_0$  to  $R_1$  and parallel to the disks plane.

In all cases, the boundary conditions correspond to no-slip conditions for  $u$  and  $w$  at the rigid walls. For the open cavity, an Ekman boundary layer flow [2] is considered at  $R_0$  and  $R_1$ . For rotor-stator cavities, the azimuthal velocity component at the boundaries conditions are  $v = 0$  on the stator,  $v = (R_m + r)/(R_m + 1)$  on the rotating disk.

### 3 Numerical model

The solution method is based on a pseudo-spectral collocation Chebyshev-Galerkin Fourier method. The approximation of any flow variable  $\Psi = (u, v, w, p)$  is derived from the following developments in truncated series, with  $-1 \leq r, z \leq 1$  and  $0 \leq \theta \leq 2\pi$ :

$$\Psi_{NMK}(r, z, \theta, t) = \sum_{p=K/2}^{K/2-1} \sum_{n,m=0}^{N,M} \hat{\Psi}_{nmp}(t) T_n(r) T_m(z) e^{ip\theta}$$

$T_n, T_m$ , are the Chebyshev polynomials and  $\hat{\Psi}_{nmp}$  the spectral coefficients.

The time scheme is based on the second-order backward differentiation formula for the diffusive term and on the Adams-Bashforth scheme for the nonlinear terms. The velocity-pressure coupling is performed with a projection algorithm [3,4]. For the cylindrical cavity (involving the axis), the numerical approximation method was modified with a dependent variables transformation ( $\tilde{V} = \bar{r}V$  and  $\tilde{p} = \bar{r}p$ ). In this case the natural boundary conditions on  $V$  and  $p$  are then assigned to be zero for  $\bar{r} = 0$ . The details of the technique are developed elsewhere in [5].

### 4 Rotor-stator cavity

The velocity fluctuations display the spatial structure of instabilities and are computed at given instants with respect to the average flow solution. The local Reynolds number is defined by  $Re_\delta = \Omega \delta r / \nu$ , where  $\delta = (\nu / \Omega)^{0.5}$  is the length scale of the rotating boundary layer.

Two kinds of axisymmetric instabilities have been computed succeeding to the stationary basic flow. The first one is a stationary instability on the Bödewadt layer, characterized by 3 pairs of circular rolls and which is observed (for the first time numerically) in the annular cavity ( $Rm = 5, L = 5$ ) for  $Re = 330$ . This stationary axisymmetric solution has closely similar characteristics to the type II instability showed in experiments by [6].

When further increasing the rotation rate the instability is time-dependent in both cases of the annular cavity and the cylindrical cavity. In the annular cavity ( $R_m = 4, L = 5$ ), for  $Re = 400$  the solution is oscillatory ( $\sigma = 4.7$ ). The axisymmetric vortices are visible along the two layers and travel following the flow as in recent experiments by [7].

In the cylindrical cavities closer to the axis, the axisymmetric instabilities appears for  $Re = 4000, L = 2$  and  $Re = 1600, L = 5$ . Differently to the annular cavity the Ekman layer over the rotating disk keeps stable. These vortices travel slowly inward in the Bödewadt layer down to near the axis which corresponds to a local Reynolds number  $Re_\delta \simeq 27$ . These results in cylindrical cavities are closely similar with experimental results of [8] for a cavity of aspect ratio  $L = 0.5$ , who was the first to observe traveling circular waves during an impulsive spin-down for  $25 \leq Re_\delta \leq 125$ . Then following the analysis of Savas [8] the instabilities are of type II in both the Bödewadt and the Ekman layers.

The rolls that progress under the form of rings in the axisymmetric solution, now constitute spirals arms. The transition to three-dimensional patterns was accelerated via "artificial" initial disturbances of general form  $\alpha \sin(p\theta)$  where  $p$  is an arbitrary number corresponding to an azimuthal wavelength and  $\alpha$  the magnitude rate (scaled with respect to the azimuthal velocity). The angle of the spiral patterns can be defined by  $\varepsilon$  as the orientation of the wavefront with respect to the azimuthal direction (it is defined positive when it is rolled up towards the axis of the disk in the rotation direction).

Spiral patterns (unmixed with circular pattern) arise in annular cavities at large distances from the axis ( $R_m = 5$ ) but also in the near axis region of cylindrical cavity where the confinement is important ( $L = 2$ ). In the case of the annular cavity, an oscillatory solution ( $\sigma = 21.4$ ) is obtained with 22 spiral arms in both layers. Inside the Ekman layer, the angle of the spiral wavefront  $\varepsilon$  decreases with  $\bar{r}$  as  $-15.3^\circ \leq \varepsilon \leq -6.9^\circ$  (Figure 1a). These 3D spiral

patterns have already been observed in experiments of the Ekman layer (see a review in Ref. [9]) that are referred as the type II instability. Close to the stationary disk the spatial structure of the vortices shows some similarities with that one on the rotating disk layer but the spiral arms form a positive angle,  $11^\circ \leq \varepsilon \leq 23^\circ$ .

Mixed annular and spiral pattern is also observed in the cylindrical cavity of large aspect ratio  $L = 5$  and in the annular cavity closer to the axis ( $R_m = 4$ ,  $L = 5$ ). For  $Re = 400$  in the annular cavity, 18 spiral arms can be counted inside the Ekman layer, with an angle  $\varepsilon$ , steeply decreases between  $R_0$  and  $R_1$ , from  $-20^\circ$  to  $-7.5^\circ$ . The number of arms modifies between the shaft at  $R_0$  and the shroud at  $R_1$  exhibiting a zone with dislocations (Figure 1b). These spiral patterns have characteristic parameters very close to the previous ones described for  $L = 5$ ,  $R_m = 5$  and are relevant to type II instability. Inside the Bödewadt layer, we observe in the annular domain 4 pairs of spiral and annular rolls. The latter develop in rings and the first develop into 18 spiral arms forming an angle  $15.6^\circ \leq \varepsilon \leq 23^\circ$  close to the shaft. The persisting axisymmetric structures interact with the spiral arms and travel inward. The coexistence of these two types of waves was described by [8] who identifies these spirals to the type I instability of Bödewadt layer.

In the cylindrical cavity ( $L = 5$ ) for  $Re = 1200$ , the temporal behavior exhibits an oscillatory regime ( $\sigma \cong 1$ ). For  $1.54 \leq \bar{\tau} \leq 5$ , the vortices transform under the form of 5 annular structures while at larger distance from the axis rolls develop under the form of 6 spiral structures forming an angle  $7^\circ \leq \varepsilon \leq 28^\circ$  and exhibit pairing phenomena (Figure 1c).

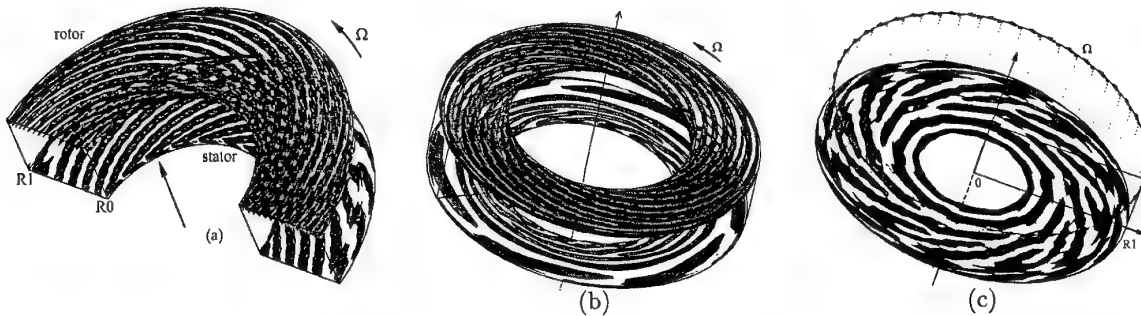


Figure 1: Three - dimensional display of iso-surfaces of the axial velocity component fluctuation. (a) Spiral patterns of the instability in annular cavity ( $R_m = 5$  and  $L = 5$ ) for  $Re = 330$ . (b) Spiral and annular patterns ( $R_m = 4$  and  $L = 5$ ) for  $Re = 400$ . (c) Spiral and annular patterns of the instability of the Bödewadt layer in the cylindrical interdisk cavity ( $L = 5$ ) for  $Re = 1600$ .

## 5 Rotating cavity submitted to a radial outflow

The base flow solution is steady and axisymmetric and organizes itself symmetrically and parallel Ekman layer flows form on the two disks with the same mass flow rate [9]. In all cases the Reynolds number is  $Re = 1750$ .

For  $C_w = 530$  the flow is axisymmetric and oscillatory ( $\sigma = 7.4$ ) (Figure 2a). We obtain good agreement between the computed value of  $\sigma$  and the frequency reported from experiments by [10]. In the numerical solution six pairs of counter-rotating axisymmetric rolls are exhibited in the Ekman layer. A good agreement with the theoretical results is obtained. The range of the characteristic parameters is relevant with the axisymmetric mode of the type II instability.

Multiple periodic solutions with different numbers of spiral arms following the periodicity of the disturbances,  $n \geq p$ ,  $p \geq \Delta R/\lambda_r$  ( $\approx 6$ ) are obtained (Figure 2b). These 3D spiral patterns have already been observed in experiments of the Ekman layer (see a review in Ref. [5]) and the characteristic parameters are in good agreement with those obtained in the relevant experiments [10]. Thus, the spiral structure of the computed rotor layer shows the same characteristics as the type II standard instability of the Ekman layer.

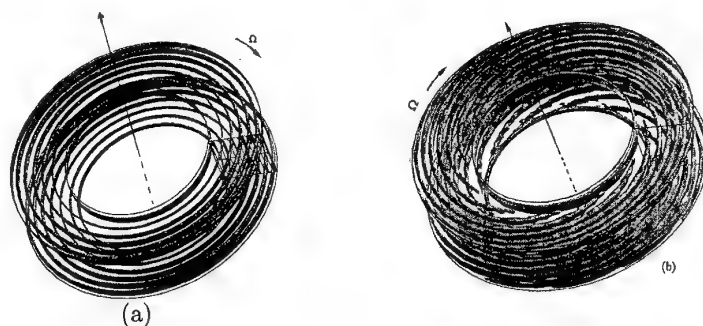


Figure 2: Three - dimensional displays of instantaneous iso-surfaces of the axial velocity component in the annular domain ( $R_m = 5$ ,  $L = 3.37$ ), for  $C_w = 530$ ,  $Re = 1750$ . (a) Circular patterns of the axisymmetric instability.(b) Spiral patterns.

## 6 Conclusion

A Fourier-Chebyshev pseudospectral method has been proposed for solving the unsteady, 3D Navier Stokes equations in rotating cavities. The main computational difficulties come from the presence of the singularity at the axis and the incompressibility constraint. We obviated the first difficulty with a transformation of the variables and the second one by means of a projection scheme. Three-dimensional simulations have shown axisymmetric and spiral instability patterns which are in good agreement with experimental studies and theoretical analyses for the type II and type I instabilities.

### Acknowledgements

*This research was supported by a research grant from DGA and CNRS. The authors gratefully acknowledge Prof. M. Schaefer, Prof. J.M. Owen and Dr. M. Wilson for fruitful discussions. The computations were carried out on Cray computer C98 on the C.N.R.S. I.D.R.I.S. Computing Center with supports from the Scientific Committee. The C.N.R.S.-D.F.G. Research Program "Numerical Simulation of Flows" and research contracts involving D.G.A. are also acknowledged.*

### 2. References

- 1 GREENSPAN, H. P. : The theory of rotating fluids. Cambridge University Press (1969).
- 2 HIDE, R. : On source-sink flows stratified in a rotating annulus. J. Fluid Mech. **32** (1968), 737-764.
- 3 HUGUES, S. & RANDRIAMAMPANINA, A. : An improved projection scheme applied to pseudospectral methods for the incompressible Navier-Stokes equations. Intl. J. Numer. Meth. Fluids **28** (1998), 501-521.
- 4 RASPO, I. : Méthodes Spectrales et de Décomposition de Domaine pour les Ecoulements Complexes Confinés en Rotation. Ph. Deg. Thesis, Université de la Méditerranée (1996)
- 5 CRESPO DEL ARCO, E., MAUBERT, P., RANDRIAMAMPANINA, A. & BONToux, P. : Spatio temporel behaviour in a rotating annulus with a source-sink flow. J. Fluid Mech. **32** (1996), 1-27.
- 6 SERRE, E. & PULICANI, J. P. : A 3D pseudospectral method for convection in a rotating cylinder. Computers and Fluids (in print).
- 7 SIRIVAT, A. : Stability experiment of flow between a stationary and rotating disk. Phys. Fluids **A3** (11) (1991), 2664-2671.
- 8 SCHOUVEILER, L., LE GAL, P., CHAUVE, M. P. & TAKEDA Y. : Spiral and Circular waves in the flow between a rotating and a stationary disk. Experiments in Fluids **26** (1999), 179-187.
- 9 SAVAS, O. : Stability of Bödewadt Flow. J. Fluid Mech. **183** (1987), 77-94.
- 10 FALLER, A. J. : Instability and transition of the disturbed flow over a rotating disc. J. Fluid Mech. **230** (1991), 245-269.
- 11 SERRE, E., HUGUES, S., CRESPO DEL ARCO, E., RANDRIAMAMPANINA, A. & BONToux, P. : Spiral and circular instability patterns in an Ekman boundary layer flow. Int. J. Heat Fluid Flows (in print).
- 12 CALDWELL, D. R. & VAN ATTA, C. W. : Characteristics of Ekman boundary layer instabilities. J. Fluid Mech. **44** (1970), 79-95.

*Addresses:* DR. SERRE E., DR. RASPO I., PROF. BONToux P., L.A.B.M., CNRS / Université Aix-Marseille II, Technopôle de Château-Gombert, 38 rue F. Joliot-Curie, 13451 Marseille cedex 20, France; email: serre1@l3m.univ-mrs.fr.

PROF. PEYRET R., Laboratoire J. A. Dieudonné, Université de Nice Sophia-Antipolis, Parc Valrose - BP 71, F - 06108 Nice Cedex 02

SOLONNIKOV, V.A.

# $L_p$ -estimates of solutions of initial-boundary value problem for generalized Stokes equations

We prove the solvability of the Cauchy-Dirichlet problem for generalized Stokes equations and obtain coercive estimates of the solution in anisotropic Sobolev spaces.

Keywords: Stokes equations, Non-newtonian fluids, Hydrodynamical potentials.

The present communication is concerned with the initial-boundary value problem

$$\frac{\partial \vec{v}}{\partial t} + A(x, t, \frac{\partial}{\partial x})\vec{v} + \nabla p = \vec{f}(x, t), \quad \nabla \cdot \vec{v} = 0, \quad x \in \Omega \subset \mathbb{R}^3, \quad t \in (0, T), \quad (1)$$

$$\vec{v}(x, 0) = \vec{v}_0(x), \quad \vec{v}(x, t)|_{x \in S} = \vec{a}(x, t), \quad (2)$$

where unknown are a vector field  $\vec{v}(x, t) = (v_1(x, t), v_2(x, t), v_3(x, t))$  and a function  $p(x, t)$ . By  $A$  we mean a matrix-formed differential second order elliptic operator with real coefficients and by  $A_0$  we mean its principal part, i.e. the sum of all terms in  $A$  containing derivatives of the second order. We assume that the matrix  $A_0(x, t, i\xi)$  is positive definite for arbitrary  $\xi \in \mathbb{R}^3$  and for arbitrary fixed  $x \in \Omega$ ,  $t \in [0, T]$ . The domain  $\Omega$  is bounded,  $S \equiv \partial\Omega$ .

When  $A = -\nabla^2 I$ , then (1) is a well known Stokes system. The system (1) arises as a result of linearization of equations of motion of non-newtonian fluids.

Our main result is the following existence theorem for the problem (1), (2).

**Theorem 1.** Assume that  $S \in C^3$ , coefficients of  $A$  are bounded and the leading coefficients are continuous in  $(x, t)$  and belong to  $W_q^1(\Omega)$ ,  $1/q < \min(1/p, (p-1)/p, 1/3) + 1/3$ ,  $p > 1$ , for all  $t \in [0, T]$ . Let the data  $\vec{f}(x, t)$ ,  $\vec{v}_0(x)$ ,  $\vec{a}(x, t)$  possess the following properties:

1.  $\vec{f} \in L_p(Q_T)$ ,  $\vec{v}_0 \in W_p^{2-2/p}(\Omega)$ ,  $\vec{a}_\tau \in W_p^{2-1/p, 1-1/2p}(\Sigma_T)$ ,  $\vec{a} \cdot \vec{n} \in L_p(0, T; W_p^{2-1/p}(S))$  where  $p > 1$ ,  $Q_T = \Omega \times (0, T)$ ,  $\Sigma_T = S \times (0, T)$ ,  $\vec{a}_\tau = \vec{a} - \vec{n}(\vec{n} \cdot \vec{a})$ ;

2. there hold the compatibility conditions  $\nabla \cdot \vec{v}_0 = 0$ ,

$$\vec{v}_0(x)|_{x \in S} = \vec{a}(x, 0), \quad \text{if } p \geq 3/2, \quad (3)$$

and the condition

$$\frac{\partial \vec{a}}{\partial t} \cdot \vec{n} = \text{div}_S \vec{A}(x, t) \quad (4)$$

where  $\vec{n}$  is the exterior normal to  $S$ ,  $\text{div}_S$  is the divergence on  $S$ ,  $\vec{A} \in L_p(0, T; W_p^{1-1/p}(S))$ . The equations  $\vec{v}_0(x) \cdot \vec{n}|_S = \vec{a}(x, 0) \cdot \vec{n}$  and (4) are understood in a weak sense, and  $\vec{v}_0(x)_\tau|_S = \vec{a}_\tau(x, 0)$  is understood as an equation between two elements of the space  $W_p^{2-3/p}(S)$ , if  $p > 3/2$ , and as the condition of the boundedness of the integral

$$I = \int_0^T \int_\Omega \int_S |\vec{a}_\tau(x, t) - \vec{v}_{0\tau}(y)|^{3/2} \frac{dtdS_x dy}{(|x-y|^2 + t)^{5/2}},$$

if  $p = 3/2$  (the function  $\vec{n}$  in the expression  $\vec{v}_{0\tau}$  is extended from  $S$  into  $\Omega$  with the preservation of class).

Then the problem (1), (2) has a unique solution  $\vec{v} \in W_p^{2,1}(Q_T)$ ,  $\nabla p \in L_p(Q_T)$ , and the solution satisfies the inequality

$$\begin{aligned} \|\vec{v}\|_{W_p^{2,1}(Q_T)}^p + \|\nabla p\|_{L_p(Q_T)}^p &\leq c \left( \|\vec{f}\|_{L_p(Q_T)}^p + \|\vec{v}_0\|_{W_p^{2-2/p}(\Omega)}^p \right. \\ &\quad \left. + \|\vec{a}_\tau\|_{W_p^{2-1/p, 1-1/(2p)}(\Sigma_T)}^p + \int_0^T \|\vec{a}(\cdot, t) \cdot \vec{n}\|_{W_p^{2-1/p}(S)}^p dt + \int_0^T \|\vec{A}(\cdot, t)\|_{W_p^{1-1/p}(S)}^p dt + I_p[\vec{a}_\tau, T] \right) \end{aligned} \quad (5)$$

where

$$I_p[\vec{a}_\tau, T] = 0, \text{ if } p > 3/2, \quad I_p[\vec{a}_\tau, T] = I, \text{ if } p = 3/2, \quad I_p[\vec{a}_\tau, T] = \frac{1}{T^{p-1/2}} \int_0^T \|\vec{a}_\tau(\cdot, t)\|_{L_p(S)}^p dt, \text{ if } p < 3/2.$$

Similar theorem holds in the two-dimensional case.

For the Stokes system Theorem 1 is proved in [1] but with somewhat stronger norm of  $\vec{a} \cdot \vec{n}$  in the right hand side of (5).

A fundamental role in the proof of this estimate is played by the analysis of the model problem for the generalized Stokes system with constant coefficients in the half-space, namely,

$$\frac{\partial \vec{u}}{\partial t} + A_0\left(\frac{\partial}{\partial x}\right)\vec{u} + \nabla q = \vec{f}(x, t), \quad \nabla \cdot \vec{u} = 0, \quad x \in \mathbb{R}^3, \quad t > 0, \quad \vec{u}(x, 0) = \vec{u}_0(x), \quad (6)$$

$$\frac{\partial \vec{v}}{\partial t} + A_0\left(\frac{\partial}{\partial x}\right)\vec{v} + \nabla p = 0, \quad \nabla \cdot \vec{v} = 0, \quad x \in \mathbb{R}_+^3, \quad t > 0, \quad \vec{v}(x, 0) = 0, \quad \vec{v}|_{x_3=0} = \vec{a}(x', t) \quad (7)$$

where  $\mathbb{R}_+^3 = \{x_3 > 0\}$ ,  $x' = (x_1, x_2)$  and  $A_0(\frac{\partial}{\partial x})$  is an elliptic operator with constant real coefficients containing only second order derivatives. The ellipticity means that the matrix  $A_0(i\xi)$  is positive definite for arbitrary  $\xi \in \mathbb{R}^3$ . As shown in [2], the solution of the problem (7), (8) can be represented in the form of the simple layer potential

$$v_k(x', t) = \sum_{m=1}^3 \int_0^t \int_{\mathbb{R}^2} T_{km}(x' - y', x_3, t - \tau) h_m(y', \tau) dy' d\tau, \quad k = 1, 2, 3,$$

$$p(x, t) = -\frac{1}{4\pi} \int_{\mathbb{R}^2} \nabla \frac{1}{\sqrt{(x' - y')^2 + x_3^2}} \cdot \vec{h}(y', t) dy' + \sum_{m=1}^3 \int_0^t \int_{\mathbb{R}^2} T'_{4m}(x' - y', x_3, t - \tau) h_m(y', \tau) dy' d\tau,$$

where  $T_{km}$  are elements of the fundamental matrix of solutions of the system (6) and  $h_k$  are linear combinations of some pseudodifferential operators applied to  $a_j$  which can be written in the form

$$\begin{aligned} h_k(x', t) &= \left( \frac{\partial}{\partial t} - \sum_{\beta=1}^2 \frac{\partial^2}{\partial x_\beta^2} \right) \left( \sum_{j=1}^2 \int_0^t \int_{\mathbb{R}^2} \Gamma(x' - y', 0, t - \tau) a_j(y', \tau) dy' d\tau \right. \\ &\quad \left. + \sum_{m=1}^3 \int_0^t \int_{\mathbb{R}^2} W_{km}(x' - y', t - \tau) a_m(y', \tau) dy' d\tau \right), \quad k = 1, 2, \\ h_3(x', t) &= \left( \frac{\partial}{\partial t} - \sum_{\beta=1}^2 \frac{\partial^2}{\partial x_\beta^2} \right) \left( \sum_{j=1}^2 \int_0^t \int_{\mathbb{R}^2} W_{3j}(x' - y', t - \tau) a_j(y', \tau) dy' d\tau \right. \\ &\quad \left. - \frac{1}{\pi} \int_{\mathbb{R}^2} \frac{a_3(y', t)}{|x' - y'|} dy' + \int_0^t \int_{\mathbb{R}^2} W_{33}(x' - y', t - \tau) a_3(y', \tau) dy' d\tau \right) \end{aligned}$$

where  $h_k = (F'L)^{-1} \hat{h}_k$ ,  $F'L$  is the Fourier-Laplace transformation with respect to  $x_1, x_2, t$ ,  $\Gamma(x, t)$  is a fundamental solution of the heat equation in  $\mathbb{R}^3$ . The functions  $T_{km}(x, t)$ ,  $T'_{4m}(x, t)$  and  $W_{km}$ ,  $k, m = 1, 2, 3$ , satisfy the inequalities

$$|D_x^j T_{km}(x, t)| \leq c(|j|)(|x|^2 + t)^{-(3+|j|)/2}, \quad |D_x^j T'_{4m}(x, t)| \leq c(|j|)(|x|^2 + t)^{-(4+|j|)/2},$$

$$|D_x^j W_{km}(x', t)| \leq c(j)(|x'|^2 + t)^{-(3+|j|)/2}, \quad k + m < 6, \quad |D_x^j W_{33}(x', t)| \leq c(j)t^{-1/2}(|x'|^2 + t)^{-(2+|j|)/2}$$

for arbitrary  $t > 0$ , and they vanish, if  $t < 0$ .

We show that if

$$\vec{a}(x_1, x_2, 0) = 0, \quad \frac{\partial a_3(x_1, x_2, t)}{\partial t} = \frac{\partial A_1}{\partial x_1} + \frac{\partial A_2}{\partial x_2} \quad (8)$$

(which should be understood in the same sense as in Theorem 1), then

$$\|\vec{v}\|_{\dot{W}_p^{2,1}(\mathbb{R}_+^3 \times (0, T))}^p + \|\nabla p\|_{L_p(\mathbb{R}_+^3 \times (0, T))}^p$$

$$\begin{aligned}
 &\leq c \left( \sum_{j=1}^2 \|h_j\|_{\dot{W}_p^{1-1/p, 1/2-1/(2p)}(\mathbb{R}^2 \times (0, T))}^p + \int_0^T \|h_3(\cdot, t)\|_{\dot{W}_p^{1-1/p}(\mathbb{R}^2)}^p dt + \frac{1}{T^{p/2-1/2}} \sum_{j=1}^2 \int_0^T \|h_j\|_{L_p(\mathbb{R}^2)}^p dt \right) \\
 &\leq c \left( \sum_{j=1}^2 \|a_j\|_{\dot{W}_p^{2-1/p, 1-1/(2p)}(\mathbb{R}^2 \times (0, T))}^p + \int_0^T \|a_3(\cdot, t)\|_{\dot{W}_p^{2-1/p}(\mathbb{R}^2)}^p dt + \int_0^T \|\vec{A}(\cdot, t)\|_{\dot{W}_p^{1-1/p}(\mathbb{R}^2)}^p dt \right. \\
 &\quad \left. + \frac{1}{T^{p-1/2}} \sum_{j=1}^2 \int_0^T \|a_j\|_{L_p(\mathbb{R}^2)}^p dt \right). \quad (9)
 \end{aligned}$$

Here the dot over  $W$  indicates that the corresponding norm is homogeneous and contains only principal terms, for instance,

$$\|\vec{u}\|_{\dot{W}_p^{2,1}(\mathbb{R}_+^3 \times (0, T))}^p = \int_0^T \left( \|D_t \vec{u}(\cdot, t)\|_{L_p(\mathbb{R}_+^3)}^p + \sum_{|j|=2} \|D_x^j \vec{u}(\cdot, t)\|_{L_p(\mathbb{R}_+^3)}^p \right) dt.$$

Inequality (9) and the estimate for the solution of the Cauchy problem (6)

$$\|\vec{u}\|_{\dot{W}_p^{2,1}(\mathbb{R}^3 \times (0, T))}^p + \|\nabla q\|_{L_p(\mathbb{R}^3 \times (0, T))}^p \leq c \left( \|\vec{f}\|_{L_p(\mathbb{R}^3 \times (0, T))}^p + \|\vec{u}_0\|_{\dot{W}_p^{2-2/p}(\mathbb{R}^3)}^p \right) \quad (10)$$

(obtained also in [3]) make it possible to prove (5) by a standard Schauder procedure (cf. [2]).

Let us say a few words about the proof of this inequality. We introduce a vector field  $\vec{w}(x, t) = \nabla \Phi(x, t)$  where  $\Phi$  is a solution of the Neumann problem

$$\nabla^2 \Phi = 0, \quad x \in \Omega, \quad \frac{\partial \Phi}{\partial n} \Big|_S = \vec{a} \cdot \vec{n}$$

The problem is solvable, since  $\int_S \vec{a} \cdot \vec{n} dS = \int_0^t d\tau \int_S \operatorname{div}_S \vec{A}(x, \tau) dS = 0$ , and, as shown in [4],

$$\|\nabla \Phi(\cdot, t)\|_{\dot{W}_p^2(\Omega)}^p + \|\Phi_t(\cdot, t)\|_{L_p(\Omega)}^p \leq c \left( \|\vec{a}(\cdot, t) \cdot \vec{n}\|_{\dot{W}_p^{2-1/p}(S)}^p + \|\vec{A}(\cdot, t)\|_{\dot{W}_p^{1-1/p}(S)}^p \right).$$

For the differences  $\vec{v}_1 = \vec{v} - \nabla \Phi$ ,  $p_1 = p + \Phi_t$  we have the problem of the same type as (1), (2) but with  $\vec{a} \cdot \vec{n} = 0$ , so we assume that this condition is satisfied by  $\vec{a}$  in (2).

Let us estimate  $\vec{v}(x, t)$  and  $p(x, t)$  in the neighbourhood of arbitrary point  $x^{(0)} \in S$  on a small time interval  $(0, t_0)$  (this restriction is easily removed). Assume that the point  $x^{(0)}$  coincides with the origin of our coordinate system and that the  $x_3$ -axis is directed along the interior normal  $\vec{n}(x^{(0)}) = \vec{n}(0)$ . Let  $x_3 = F(x_1, x_2) \equiv F(x')$  be the equation of  $S$  in  $d$ -neighbourhood of the origin, and let  $\psi_\lambda(y)$ ,  $\lambda \in (0, d/2)$ , be a smooth cut-off function equal to one for  $|y| \leq \lambda$ , to zero for  $|y| > 2\lambda$  and satisfying the inequalities  $0 \leq \psi_\lambda(y) \leq 1$ ,  $|D^j \psi_\lambda(y)| \leq c(j)\lambda^{-|j|}$ . We make the change of variables near the origin, according to the formula  $y' = x'$ ,  $y_3 = x_3 - F(x')$ , and we introduce the functions  $\vec{u} = \psi_\lambda \vec{v}$ ,  $q = \psi_\lambda p$ . They satisfy the relations

$$\frac{\partial \vec{u}}{\partial t} + A_{00} \left( \frac{\partial}{\partial y} \right) \vec{u} + \nabla q = \vec{f} \psi_\lambda + (\nabla - \tilde{\nabla}) q + (A_{00} - \tilde{A}) \vec{u} + \vec{f}_1(y, t) \equiv \vec{h}(y, t), \quad (11)$$

$$\nabla \cdot \vec{u} = (\nabla - \tilde{\nabla}) \cdot \vec{u} + \vec{v} \cdot \tilde{\nabla} \psi_\lambda \equiv g(y, t),$$

$$\vec{u}|_{t=0} = \vec{v}_0 \psi_\lambda, \quad \vec{u}|_{y_3=0} = \vec{a} \psi_\lambda \quad (12)$$

where  $\tilde{\nabla}$  is a transformed gradient:  $\tilde{\nabla} = (\frac{\partial}{\partial y_1} - F'_{y_1} \frac{\partial}{\partial y_3}, \frac{\partial}{\partial y_2} - F'_{y_2} \frac{\partial}{\partial y_3}, \frac{\partial}{\partial y_3})$ ,  $\tilde{A}$  is a transformed operator  $A$ ,  $A_{00}(\frac{\partial}{\partial y}) = A_0(0, 0, \frac{\partial}{\partial y}) = \tilde{A}_0(0, 0, \frac{\partial}{\partial y})$  and  $\vec{f}_1 = \tilde{A}(\vec{v} \psi_\lambda) - \psi_\lambda \tilde{A} \vec{v} + p \tilde{\nabla} \psi_\lambda$ . We extend  $\vec{u}$  and  $q$  by zero into the domain  $|y| > 2\lambda$ ,  $y_3 > 0$  and we consider (11), (12) as equations in  $\mathbb{R}_+^3$ .

Further, we introduce the vector field  $\vec{w}_1(y, t) = \nabla \Psi(y, t)$  where  $\Psi$  is a solution of the Neumann problem

$$\nabla^2 \Psi = g(y, t), \quad y \in \mathbb{R}_+^3, \quad \frac{\partial \Psi}{\partial y_3} \Big|_{y_3=0} = \psi_\lambda a_3(y', t).$$

Then  $\vec{u}_1 = \vec{u} - \nabla \Psi$ ,  $q_1 = q + \Psi_t$  satisfy the relations

$$\frac{\partial \vec{u}_1}{\partial t} + A_{00} \left( \frac{\partial}{\partial y} \right) \vec{u}_1 + \nabla q_1 = \vec{h}_1, \quad \nabla \cdot \vec{u} = 0,$$

$$\vec{u}_1|_{t=0} = \vec{v}_0 \zeta_\lambda - \nabla \Psi(x, 0) \equiv \vec{v}_0, \quad \vec{u}|_{y_3=0} = \vec{a} \psi_\lambda - \nabla \Psi(y, t)|_{y_3=0} \equiv \vec{b}(y', t)$$

where  $\vec{h}_1 = \vec{h} - A_{00}(\frac{\partial}{\partial y})\vec{w}_1$ . It is easily seen that  $b_3 = 0$ . The function  $\Psi$  is given by

$$\Psi(y, t) = - \int_{\mathbb{R}_+^3} \frac{\partial N(y, z)}{\partial z_3} \nabla F(z') \cdot \vec{u}(z, t) dz + \int_{\mathbb{R}_+^3} N(y, z) \vec{v} \cdot \vec{\nabla} \psi_\lambda dz + \int_{\mathbb{R}^2} N(y, z', 0) (a_3 - \nabla F \cdot \vec{a}) \psi_\lambda(z', 0) dz'$$

where  $N$  is the Green function of the Neumann problem in  $\mathbb{R}_+^3$ . The function  $D_t(a_3 - \nabla F \cdot \vec{a})$  is representable in the form (8), so we may estimate  $\vec{w}_1$  with the help of the Calderon-Zygmund theorem and of the well known bounds for  $L_p$ -norms of the second derivatives of the solution of the Neumann problem in  $\mathbb{R}_+^3$ . The functions  $\vec{u}_1, q_1$  can be estimated by inequalities (9), (10). If  $\lambda$  and  $t_0$  are small enough, standard calculations lead to

$$\begin{aligned} \|\vec{v}\|_{W_p^{2,1}(Q_{\lambda,t})}^p + \|\nabla p\|_{L_p(Q_{\lambda,t})}^p &\leq c \left( \|\vec{f}\|_{L_p(Q_{2\lambda,t})}^p + \|\vec{v}_0\|_{W_p^{2-2/p}(\Omega_{2\lambda})}^p + \|\vec{a}\|_{W_p^{2-1/p, 1-1/(2p)}(\Sigma_{2\lambda,t})}^p + I_p[\psi_\lambda \vec{a}, t] \right) \\ &\quad + c(\lambda) \left( \|\nabla \vec{v}\|_{L_p(Q_{2\lambda,t})}^p + \|\vec{v}\|_{L_p(Q_{2\lambda,t})}^p + \|p\|_{L_p(Q_{2\lambda,t})}^p \right), \quad \forall t \in (0, t_0), \end{aligned}$$

where  $\Omega_\lambda = \{x \in \Omega : |x| < \lambda\}$ ,  $Q_{\lambda,t} = \Omega_\lambda \times (0, t)$ ,  $S_\lambda = \partial\Omega_\lambda \cap S$ ,  $\Sigma_{\lambda,t} = S_\lambda \times (0, t)$ . The same estimate can be obtained in the neighbourhood of arbitrary point  $\bar{\Omega}$ , so we arrive at

$$\begin{aligned} \|\vec{v}\|_{W_p^{2,1}(Q_t)}^p + \|\nabla p\|_{L_p(Q_t)}^p &\leq c \left( \|\vec{f}\|_{L_p(Q_t)}^p + \|\vec{v}_0\|_{W_p^{2-2/p}(\Omega)}^p + \|\vec{a}\|_{W_p^{2-1/p, 1-1/(2p)}(\Sigma_t)}^p + I_p[\vec{a}, t] \right) \\ &\quad + c \left( \|\nabla \vec{v}\|_{L_p(Q_t)}^p + \|\vec{v}\|_{L_p(Q_t)}^p + \|p\|_{L_p(Q_t)}^p \right), \quad \forall t \in (0, t_0). \end{aligned}$$

We consider  $p$  as a solution of the Neumann problem

$$\nabla^2 p = -\nabla \cdot (A\vec{v} - \vec{f}), \quad \frac{\partial p}{\partial n} \Big|_S = -\vec{n} \cdot (A\vec{v} - \vec{f})|_S$$

which allows us to show that

$$\|p(\cdot, t)\|_{L_p(\Omega)} \leq c \left( \|\nabla \vec{v}\|_{L_p(\Omega)} + \|\nabla \vec{v}\|_{L_p(S)} + \|\vec{v}\|_{L_p(\Omega)} + \|\vec{f}\|_{L_p(\Omega)} \right), \quad \forall t \in (0, t_0)$$

(cf [1]). The norms of  $\nabla \vec{v}$  and  $\vec{v}$  can be estimated with the help of interpolation inequalities and of the Gronwall lemma which leads to (5).

The work was done at the Mathematical center (CMAF) of the University of Lisbon whose hospitality is gratefully acknowledged.

## 1. References

- 1 SOLONNIKOV, V.A.: Estimates of solutions of nonstationary linearized system of the Navier- Stokes equations. Trudy Math. Inst. Steklov, **73** (1964), 213-317.
- 2 SOLONNIKOV, V.A.: Initial-boundary value problem for generalized Stokes equations. Mathematica Bohemica, **1** (2001), to appear.
- 3 LADYZHENSKAYA, O.A., SEREGIN, G.A.: Coercive estimates for solutions of linearizations of modified Navier-Stokes equations. Dokl. Acad. Nauk, **370** (2000), 738-740.
- 4 SOLONNIKOV, V.A.: Estimates of solutions of nonstationary Navier-Stokes equations. Zap. Nauchn. Semin. L.O.M.I., **38** (1973), 153-231.

Address: Prof. Dr. V.A.Solonnikov, V.A.Steklov Math. Inst., Fontanka 27, 191011, S.Petersburg, Russia.



SCHIRMACHER, ROLF

## The Use of Active Vibration Control for the Reduction of ICE Interior Noise

*First generation ICE high speed trains show a disturbing low frequency noise of about 100 Hz audible inside the coaches. It is excited by unround wheels, propagates via the bogie to the body of the coach and is finally radiated as airborne sound.*

*A mixed concept of active suspension at the bogie in connection with adaptive residual noise minimisation inside the passenger compartment is successfully applied to the problem. A prototype system utilizing piezo ceramic actuators was installed and tested on the rolling rig of Deutsche Bahn in Munich. The noise level reductions at single harmonics were more than 12 dB averaged over the whole compartment and more than 20 dB at single seats. Measurement results and practical experiences with the system are reported.*

### 1. Introduction

Low frequency interior noise poses an important problem for modern high speed railcars. One example for such a railcar is the German high speed train ICE. With the first generation of this high speed train, ICE 1, DB (Deutsche Bahn, German Rail) is operating high speed rail connections with up to 280 km/h. The intermediate passenger cars of this train, equipped with bogie type MD 530, exhibit low frequency vibrations at about 100 Hz. These vibrations are excited by unround wheels (with the fourth harmonic of the wheel rotational frequency being at the 100 Hz range at about 200 km/h and the third at about 250 km/h, both typical operational speeds for the German high speed rail network) and the sleepers below the rail (the sleeper passing frequency of 100 Hz corresponds to a speed of 216 km/h).

The bogie constitutes a double elastic suspension, set up by a primary suspension directly on top of the axle bearings and a secondary suspension below the bogie bolster. Figure 1 shows a drawing of the Bogie MD 530. The secondary suspension exhibits spring resonances at the 100 Hz region, but there are also further resonances at this frequency range at different parts and subsystems of the railcar.

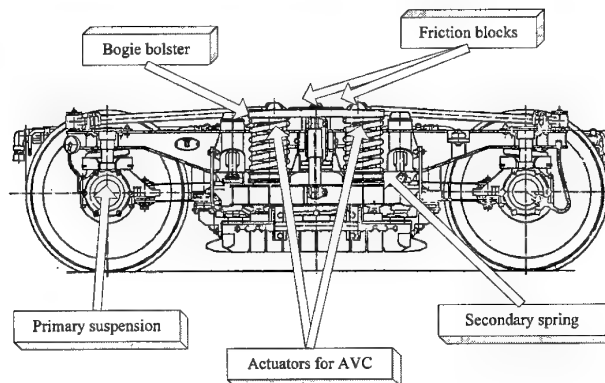


Figure 1: Bogie MD 530 as used at the ICE 1 passenger coaches. The car body is loaded on top the friction blocks.

The car body itself is a large (26.4 m long, up to 3 m wide) aluminium structure with a high modal density of structural as well as acoustical cavity modes. The vibration enters the body through the bogie-body interface and is well predictable because of the harmonic excitation.

### 2. Active Vibration Control Concept

As the railcar body is supported by only two friction blocks on each side of the bogie and there are no clearly visible flanking transmission paths (e.g., no large dampers in parallel to the secondary suspension), it was decided to set up an active vibration control system which actuates at the bogie-body-interface [1,2]. By this, four point forces

should be sufficient for an active control system suppressing the vertical vibration. The horizontal degrees of freedom were assumed to be of minor importance, as the secondary spring should be softer in the off-axis plane. Preliminary experiments and field measurement data (high speed train rides) gave a need for an active control force of about 200 N at 90 Hz for a predominant wheel harmonic. This was set as a design target for the actuator systems. By using actuators at this interface, the necessary control forces are comparatively low because the whole passive suspension is still used to reduce them.

As there surely are flanking transmission paths in parallel to the friction blocks, it was decided to use error microphones distributed inside the car as control inputs. By this, some aspects of Active Structural Acoustic Control (ASAC, minimisation of the radiated air-borne sound instead of some structural quantity) are combined with aspects of active vibration isolation (cutting the power flow at some prominent points). This concept turned out later on to be a key element of the success of the system as it really shows that acoustic minimisation does not necessarily imply the minimisation of vibrational quantities.

To generate the compensation signal for the actuators, a feed-forward scheme utilizing a revolution counter was used. By this, any feed-back problems of the actuator output to the control system reference signal are eliminated. This controller design lead to a very robust active system.

### 3. Actuator System

The most critical system components are the actuators. This is a very typical situation for active noise and vibration control. For the active control of the 100 Hz vibration, about 200 N at 90 Hz are needed on each side of the bogie according to preliminary investigations. The actuators also have to fit into the bogie design with its space restrictions. As a result, piezoceramic driven inertial mass actuators were developed. These actuators fit into the free space of the inner secondary suspension springs even under all driving conditions of the bogie. Figure 2 shows a sketch of such an actuator system.

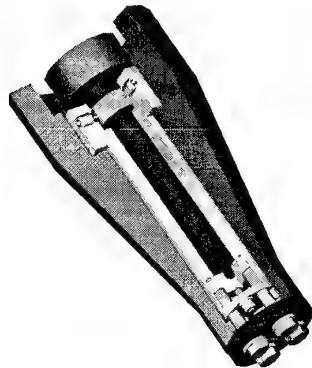


Figure 2: Inertial mass actuator system as used for the active control of the ICE 1 vibrations. 11 kg lead mass suspended from a piezoceramic drive unit with a maximum amplitude of  $140 \mu\text{m}_{\text{pp}}$  and fitting into the inner secondary suspension spring.

The inertial mass is a lead hull mass carried by an inner steel element. The drive unit is a stacked high voltage piezoceramic actuator giving a maximum displacement of  $140 \mu\text{m}_{\text{pp}}$ . Two of these elements, each hanging inside the inner secondary suspension springs, give the necessary dynamic force at each side of the bogie bolster. The piezos were driven by a 100 W high voltage amplifier each.

As the actuator systems could be made out of commercially available key components, these actuators were very cost-effective for the experiments. Further on, much experience on the practical applicability of 'new' actuators was gained throughout the development as well as the experimental work.

### 4. Experimental Results

A bogie MD 530 was equipped with the active vibration control system as described above and mounted below an ICE 1 intermediate car. The whole car was tested on the rolling rig of DB in Munich. This research facility allows test drives up to more than 500 km/h, curve simulations by car tilting and low frequency (up to about 20 Hz) track simulation of previously measured existing railway lines. By this, the sleeper passing frequency is not modelled on

this test stand, but the whole range of vehicle ride dynamics is modeled by the facility.

Tests were performed with different wheelsets showing different unroundnesses of the wheels, for different speeds, for run-ups, for different simulated lines and with and without tilting for curve simulation. At all these various conditions, interior noise reductions in the 80 to 120 Hz range (the working range of the control system) were found. With constant driving conditions, for single wheel harmonics, spatially averaged reductions were more than 12 dB with single seat reductions higher than 20 dB. Figure 3 shows some typical results in the upper half.

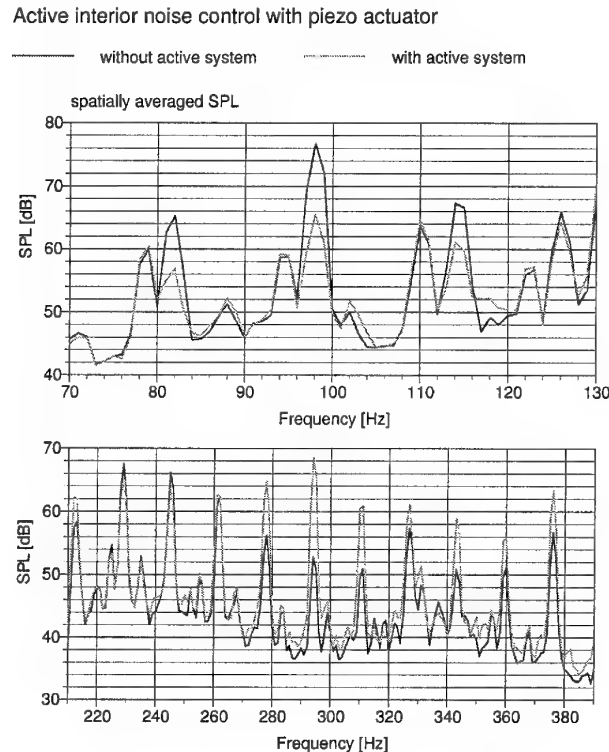


Figure 3: Spatially averaged sound pressure level inside the ICE 1 passenger car at a rolling rig test drive with 160 km/h, extremely unround wheelset. Top: working range of the active control system showing about 12 dB spatially averaged reduction at the fifth harmonic at 98 Hz while also operating on the fourth and sixth harmonic (the system was not set up to work on the second wheelset with a slightly higher diameter and thus lower frequencies as well, the 95/110 Hz harmonics are excited by this second wheelset). Bottom: Frequency range of the second harmonic of the working frequencies, showing sound pressure level increases due to piezo nonlinearities.

As shown on the lower half of figure 3, the actuator drive units are not usable that way for a practical system. Due to the nonlinearities of the piezoceramic based on the varying dynamic load during the cycle as well as on the hysteresis, the ceramics produce quite significant higher harmonic distortion. As the actuator system was driven below its mechanical resonance, these distortions are further amplified in terms of dynamic force compared to the displacement nonlinearities typically stated for piezoceramic materials. As a result, the sound pressure level did for some measurements notably increase at the higher harmonics frequency range, mostly for the second harmonic of the active control frequencies.

This drawback does not only occur with piezoceramics but was also shown with magnetostrictive materials at test stand experiments in a comparable environment [4]. Although much research is focused on that topic, including modelling and real-time control strategies for the materials, no practical solution has been found so far.

Nevertheless the validity of the concept and the feasibility of active interior noise control for high speed railcars was demonstrated by the experiments as high reductions of the interior sound pressure levels were reached.

Figure 4 demonstrates the validity of the concept regarding the mixture of active suspension and interior noise control. In that figure, the sound pressure level distribution inside the car is shown with the system switched on and off for a 200 km/h test ride. As can be seen, the sound pressure level is reduced through the whole interior space by the application of only four dynamic point forces. It should be noted that the horizontal vibration levels at the bogie bolster as well as some vertical vibration levels, e.g., at the floor of the passenger compartment increased due

to the operation of the active system. This clearly demonstrates the benefits of the concept of error microphones and compensation forces.

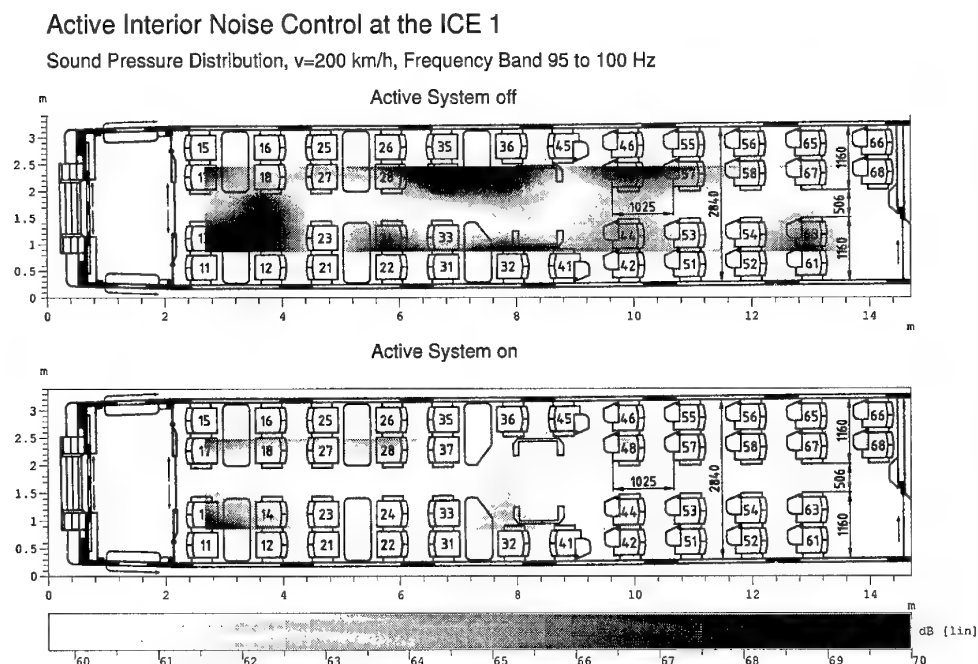


Figure 4: Sound pressure distribution in the 95 to 100 Hz band inside the ICE 1 passenger compartment, at a test ride with 200 km/h on the rolling rig. Active system off (top) and on (bottom) showing the global reduction through the whole interior space.

## 5. Conclusion

The feasibility of an active low frequency interior noise reduction system for a high speed train was demonstrated by rolling rig measurements. A concept of sound minimisation by control forces showed up to be essential for the success of the system. The piezoceramic actuators showed to be unusable for practical applications due to nonlinearities. Nevertheless, significant noise level reductions in the working range of the system were demonstrated.

## Acknowledgements

*The work presented in this paper was sponsored by Deutsche Bahn AG, Research Center Munich. The rolling rig tests would not have been possible without the support from my colleagues at Müller-BBM and the whole team at the rolling rig of Deutsche Bahn AG in Munich.*

## 6. References

- 1 SCHIRMACHER, R., HÖLZL, G., REDMANN, M. AND SCHEUREN, J.: Active Noise and Vibration Control for a High Speed Railcar: A Case Study. Proc. ACTIVE 97, 557-564.
- 2 SCHIRMACHER, R., HÖLZL, G. AND SCHEUREN, J.: Untersuchung zur aktiven Minderung tieffrequenter Vibrationen am ICE 1, DAGA 98, 210-211.
- 3 SCHIRMACHER, R., HÖLZL, G. AND SCHEUREN, J.: Active noise and vibration control for a high-speed railcar: Rolling rig measurements, Forum Acusticum 1999 Paper 4aSAb2, ACUSTICA - acta acustica 85 (1999), Suppl. 1, S316.
- 4 SCHIRMACHER, R.: Piezo actuators for active vibration control and isolation – principles and practical experiences, Forum Acusticum 1999 Paper 5pSAb7, ACUSTICA - acta acustica 85 (1999), Suppl. 1, S469.

Address: DR. ROLF SCHIRMACHER, Müller-BBM GmbH, Robert-Koch-Strasse 11, 82152 Planegg, Germany.

STÖBENER, U.; GAUL, L.

## Active Vibration and Noise Control by Hybrid Active Acoustic Panels

*In the present paper a hybrid passive and active treatment for vibration and noise reduction of plate type structures is proposed. The treatment is manufactured as sandwich structure and is called Hybrid Active Acoustic Panel. The passive component is used to reduce the vibration and sound radiation for high frequencies whereas the active part of the system is designed for the low frequency range. By selecting the thickness of the passive damping layer a certain frequency limit is defined, which divides the high and low frequency range. The actuator and sensor layout of the active component is evaluated by using the mode shapes of the low frequency range. According to the evaluated layout a Hybrid Active Acoustic Panel is manufactured and experimentally tested. The experimental results validate the proposed concept.*

### 1. Introduction

Different concepts of hybrid passive and active damping treatments have been developed and proposed, such as the Electro Mechanical Surface Damping [2,3] and the Active Constrained Layer Damping [1]. Most of these hybrid damping treatments are designed to enhance the viscous damping by increasing the shear or compressional deformation of the passive layer or to create additional energy dissipation in electric shunts. The Hybrid Active Acoustic Panel, which has been first proposed in [4], is not designed to improve passive energy dissipation but to combine viscous damping effects and active control moments.

A hybrid system for acoustic applications is proposed by VEERAMANI AND WERELEY in [6]. Their hybrid system is developed in order to reduce the sound transmission between a sound source and a receiving chamber and is based on a Kevlar-epoxy composite plate. The used damping layer (3M Scotchdamp ISD-112) is embedded within eight Kevlar plies and one PZT actuator is bonded to one surface of the plate. In contrast to this hybrid treatment the Hybrid Active Acoustic Panel is designed to reduce the sound radiation of an existing housing.

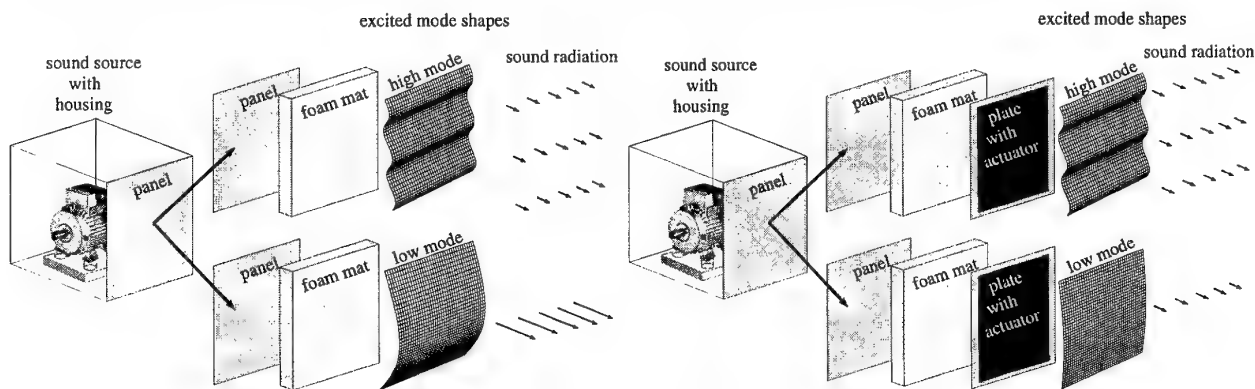


Figure 1: Concept of the Hybrid Active Acoustic Panel

Passive damping measures for the reduction of noise radiation from vibrating structures such as damping coatings or absorbers are successfully used especially for the high frequency range. On the left hand side of Fig. 1 the housing of a sound source is depicted. The housing is covered by a foam mat. The vibration of the housing causes deformations of the foam mat and due to the viscoelastic material behaviour damping takes place. The dissipated energy reduces the structural vibration and the related sound radiation as well. The reduction in the low frequency range is more difficult because the increase in space and weight is significant. Compared with passive measures, active vibration control efficiently reduces sound radiation for low frequencies by lightweight actuators. On the right hand side of Fig. 1 the foam mat and an active plate is bonded to the outer side of the housing. The remaining vibration in the low frequency range is now counterbalanced by the bending moments generated by the actuators. Therefore a combination of passive damping layers and active vibration control leads to a broadband reduction of sound radiation by a minimum of additional mass with respect to the housing.

## 2. Selection of the passive layer thickness

The Hybrid Active Acoustic Panel consists of a foam mat and a thin aluminum plate, as shown in Fig. 1 and 2. For high sound absorption coefficients the foam is made of a melamin resin. The coefficients are listed in Table 1 for two different thicknesses and five different frequencies. These data are adopted from the manufacture data sheet and are defined by DIN 5212. It can be seen that the absorption increases with higher frequencies.

frequency	125 Hz	250 Hz	500 Hz	1000 Hz	2000 Hz	4000 Hz
sound absorption coefficients (10mm mat)	0.06	0.08	0.10	0.17	0.31	0.38
sound absorption coefficients (20mm mat)	0.07	0.09	0.16	0.30	0.54	0.76

Table 1: Sound absorption coefficient of the foam mat

The selection between the two thicknesses depends on the sound absorption coefficients and the number of modes in the low frequency range. Regarding the 10 mm foam mat an appreciable decrease of radiated sound is achieved for frequencies higher than 1000 Hz. Approximately the same order of sound reduction can be obtained by using the 20 mm thick foam mat for frequencies starting from 500 Hz. Therefore the frequency limit mentioned in the abstract of this paper is 1000 Hz for the thin mat and 500 Hz for the thick plate.

The number of eigenfrequencies and the eigenfrequencies  $\omega_i$  of the base plate with free boundary conditions can be calculated analytically by the equation

$$\omega_i = \frac{\lambda_i \pi^2 t_p}{a^2} \sqrt{\frac{E}{12\rho(1-\nu^2)}}, \quad (1)$$

where  $\rho$  is the plate density,  $E$  is the Youngs modulus,  $\nu$  is the Poisson ratio and  $t_p$  is the thickness,  $a$  is the length of the plate.  $\lambda_i$  is a factor which depends on the nodal pattern and the boundary conditions as formulated, e.g., by WARBURTON in [7]. For  $\rho = 2700 \text{ kg/m}^3$ ,  $E = 70000 \text{ N/mm}^2$ ,  $t_p = 10 \text{ mm}$ ,  $a = 500 \text{ mm}$  the number of eigenfrequencies between 0 and 500 Hz is 2 whereas between 0 and 1000 Hz 7 eigenfrequencies are calculated. Taking into account that the control of 7 modes is possible and a minimum of mass and volume is desired for the hybrid panel the 10 mm foam mat is chosen for further investigations.

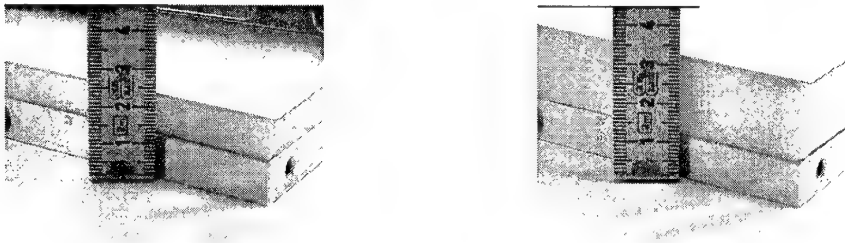


Figure 2: Base plate, 10 mm and 20 mm foam mats and aluminum plate

## 3. Layout of the active plate

Since a modal controller is used for the active vibration reduction the mode shapes of the Hybrid Active Acoustic Panel are required. The panel is a multilayer structure with nonlinear and nonisotropic properties, therefore the determination of the mode shapes is difficult. In a first step the modes of the base plate are calculated. For the calculation Kirchhoff plate theory and the formulas of WARBURTON [7] are used, where the characteristic beam functions,

$$\begin{aligned} \phi_1(x) &= 1 \text{ for } m = 0, \quad \phi_1(x) = 1 - \frac{2x}{a} \text{ for } m = 1, \\ \phi_1(x) &= \cos \gamma_1 \left( \frac{x}{a} - \frac{1}{2} \right) + k_1 \cosh \gamma_1 \left( \frac{x}{a} - \frac{1}{2} \right) \text{ for } m = 2, 4, 6, \dots, \\ \phi_1(x) &= \sin \gamma_2 \left( \frac{x}{a} - \frac{1}{2} \right) + k_2 \sinh \gamma_2 \left( \frac{x}{a} - \frac{1}{2} \right) \text{ for } m = 3, 5, 7, \dots, \\ \text{with } k_1 &= -\frac{\sin 0.5\gamma_1}{\sinh 0.5\gamma_1}, \quad \tan 0.5\gamma_1 + \tanh 0.5\gamma_1 = 0, \quad k_2 = \frac{\sin 0.5\gamma_2}{\sinh 0.5\gamma_2}, \quad \tan 0.5\gamma_2 - \tanh 0.5\gamma_2 = 0, \end{aligned} \quad (2)$$

approximate the shapes. The same functions are used for the  $y$  direction. The results of the calculation are shown in Fig. 3 in the upper row. In a second step the mode shapes of the base plate are evaluated by an experimental modal analysis. These results are listed below the calculated modes in Fig. 3. Obviously the calculation and the experiment leads to the same modes.

Since the bending stiffness of the base plate is much higher than the stiffness of the foam mat and the active plate it is assumed that the mode shapes of the hybrid panel are dominated by the shapes of the base plate. In order to validate this assumption a modal analysis is carried out for the coupled hybrid structure. It has to be noticed that the results of this analysis have to be discussed carefully since modal analysis is based on linear theory. Therefore the results, which are depicted in the lower row of Fig. 3, have to be interpreted in the context of the results for the base plate and under the mentioned assumption. Only the modes (2,0), (2,1), (3,0) and (0,2) are identified for the hybrid panel whereas the modes (1,1) and (1,2) do not occur.

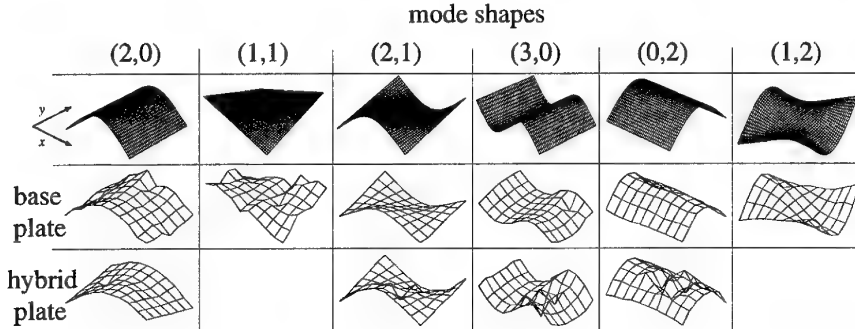


Figure 3: Calculated and experimentally evaluated mode shapes

Using the information of the identified modes the PVDF sensors are placed at the locations of maximum strain which correspond to the locations of maximum curvature. The sensors are arranged in an array and their outputs are assembled by a modal sensor matrix in order to filter modal displacements or velocities respectively. Therefore the condition of the modal sensor matrix has to be taken into consideration for the sensor placement. Details of the sensor placement are discussed in [5]. The determined sensor positions are indicated by the shaded areas in Fig. 4.

The creation of the actuator layout is based on the idea to represent the actuator function by bending moments along their edges. Therefore the actuator edges have to be adapted to the nodal lines of the mode shapes. The actuator design is also discussed in [5]. According to the mentioned design rules the layout shown in Fig. 4 has been created. The white arrows indicate the direction of induced strain.

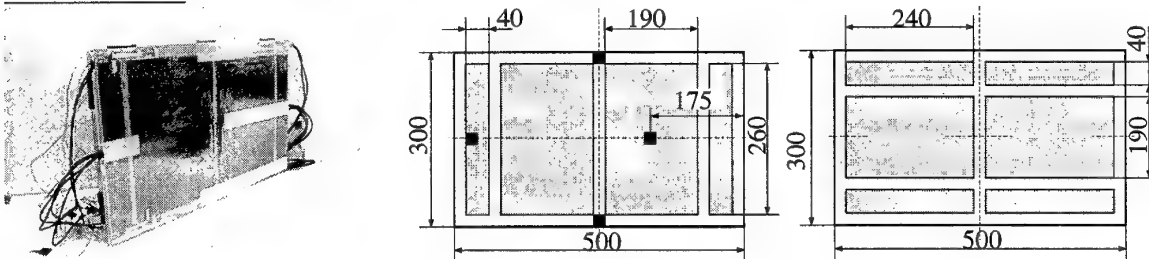


Figure 4: Test panel

#### 4. Experimental test of the panel

For experimental investigations the manufactured Hybrid Active Acoustic Panel is clamped into a wooden frame which is a part of an acoustic enclosure. To obtain free boundary conditions rubber strips are used as spacers between the frame and the base plate of the test structure. At the outer side of the acoustic enclosure an electromagnetic shaker is connected to the base plate. Inside the enclosure three microphones are positioned to sense the radiated sound pressure. Microphone 1 is located at the center of the upper edge of the panel, microphone 2 is placed at the right upper corner of the panel and microphone 3 is positioned at the center of the panel. The distance of all microphones to the surface of the panel is 20 cm. Swept sinusoidal excitation of 3 Newton amplitude is used to generate the vibrations and to acquire the radiated sound pressure for a frequency range from 0 to 800 Hz. This measurement is done for the base plate without coating, the hybrid panel without active control and the hybrid panel with active vibration control. The results are plotted in Fig. 5 and it can be seen that a significant reduction

is achieved by the passive layer. The active component enhances the reduction especially for the mode (3,0) since the chosen actuator layout enables a high controllability for this mode.

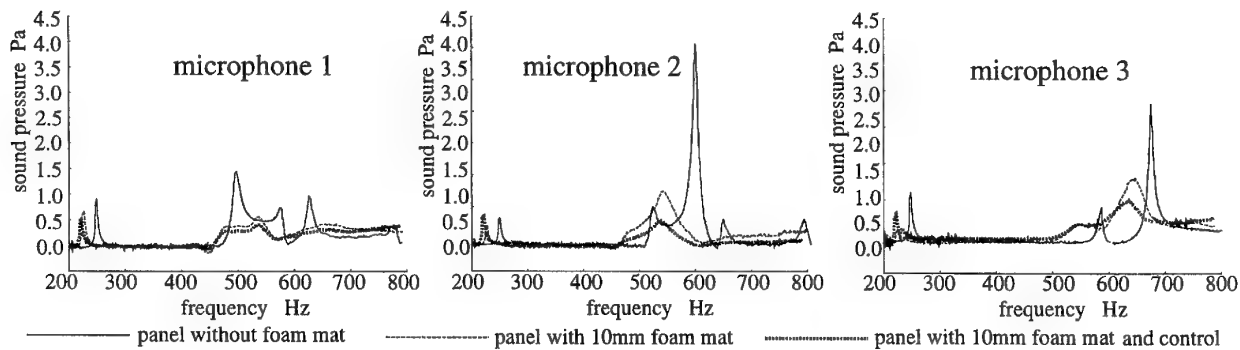


Figure 5: Sound pressure for the panel with and without control

By using the same experimental setup but replacing the microphones by a sound intensity probe the radiated sound power is acquired. The results are shown in the bar charts of Fig. 6. Obviously the main reduction is achieved by the passive layer and the active component generates an additional decrease of the sound power. The great advantage of the active component is its variability. By modifying the controller the radiation characteristics can be changed in a preselected way, e.g., to suppress dominant harmonics.

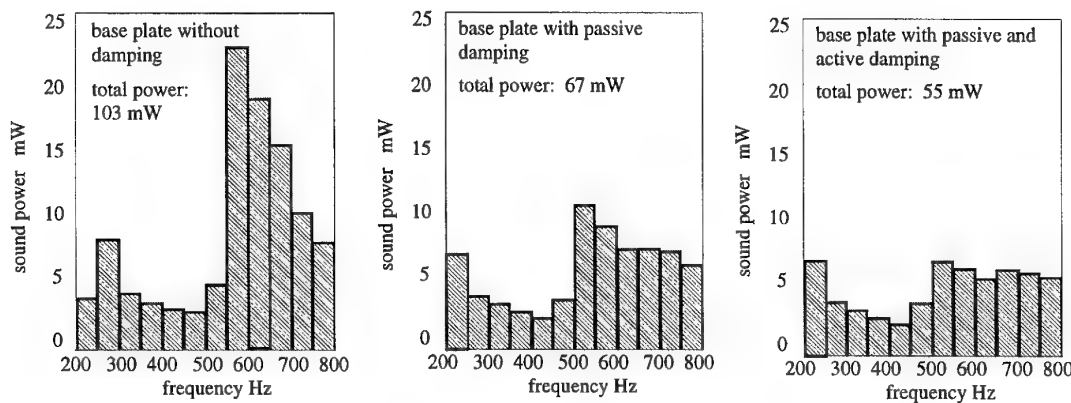


Figure 6: Sound power for the panel with and without control

## 5. References

- 1 BAZ, A. ; J. RO: Vibration Control of Plates with Active Constrained Layer Damping. *Journal of Smart Materials and Structures*, **Vol. 5** (1996), 272–280.
- 2 GHONEIM, H.: Bending and Twisting Vibration Control of a Cantilever Plate via Electromechanical Surface Damping. *Proceedings of Smart Structures and Materials Conference*, edited by C. Johnson, **Vol. SPIE-2445** (1995), 28–39.
- 3 HAGOOD, N. ; A. VON FLOTTOW: Damping of Structural Vibrations with Piezoelectric Materials and Passive Electric Networks. *Journal of Sound and Vibration*, **Vol. 146** (1991), 243–268.
- 4 STÖBENER, U. ; GAUL, L.: Control of Sound Radiation by Semi Active Acoustic Panels. *Proceedings of ACTIVE 99. Institute of Noise Control Engineering of the USA Inc.* **Vol. 1** (1999), 457–468.
- 5 STÖBENER, U. ; GAUL, L.: Sensor and Actuator Design Methods in Active Vibration Control for Distributed Parameter Structures. *Advanced Course of International Center for Mechanical Science (CISM) in Udine, June 19-23, 2000, Smart Structures: Theory and Applications*, to appear 2001
- 6 VEERAMANI, S. ; WERELEY, N.: Hybride Passive / Active Damping for Robust Multivariable Acoustic Control in Composite Plates. *Journal of Smart Structures and Materials*, **Vol. SPIE-2717** (1996), 374–387.
- 7 WARBURTON, G.B.: The Vibration of Rectangular Plates. *Proceedings of the Institution of Mechanical Engineers*, **Vol. 168** (1954), 371–384

Address: DIPL.-ING. UWE STÖBENER, Universität Stuttgart, Institut A für Mechanik  
 Allmandring 5 B, 70550 Stuttgart, Germany, e-mail: U.Stöbener@mecha.uni-stuttgart.de  
 PROF. DR.-ING. HABIL. LOTHAR GAUL, Universität Stuttgart, Institut A für Mechanik  
 Pfaffenwaldring 9, 70550 Stuttgart, Germany, e-mail: L.Gaul@mecha.uni-stuttgart.de



ALLGAIER, R.; GAUL, L.; KEIPER, W.; WILLNER, K.

## Friction Induced Vibrations and Mode Lock-In

*In this paper the phenomenon of mode lock-in is investigated, which occurs in structures where vibrations are excited by frictional contact forces. In general "mode lock-in" denotes the coupling of substructure modes which form the vibratory response of the assembly. To study this effect an experimental beam-disk set-up was built and a corresponding FE model was generated. A friction model is formulated which incorporates the surface roughness by a statistical distribution. The time-stepping solutions of the FE model were checked for system resonances and compared to the experimentally observed mode lock-in.*

### 1. Introduction

Friction-induced oscillations occur in many physical systems. A system of great practical importance is, e.g. the brake. Disc brakes can develop large sustained oscillations which are heard as "brake squeal". At present, no comprehensive model for the many phenomena of brake squeal exists. This paper presents a novel approach to the numerical modelling of brake squeal as a friction-induced oscillation. As a highly simplified physical system the beam-on-disc device is used. This apparatus has the advantage of reduced complexity by well-defined components and boundary conditions; well-defined confined contact area and a simple friction couple. However the dynamics of the beam-on-disc device has characteristic features of more complex systems. Friction in the plane of the disc leads to large out-of-plane oscillations of the disc; modes of the components "lock" into new "assembly modes" when coupled through a friction interface. The beam-on-disc device is therefore a well-suited test object for the development of numerical representations of contact laws.

### 2. Mode lock-in

The investigated system consists of two components (beam and disc) whose dynamics (natural frequencies and vibration modes) are well understood, when they are considered separately. The assembly, as coupled through a friction interface, has somewhat different system frequencies and modes. Non-linear friction coupling causes particular modes of the components to lock into one another. When mode lock-in takes place, a special base frequency with all of its higher harmonics can be observed in the resonant-like system response [1]. For example, when the second bending mode of the free beam and the (0,3) mode of the free disc are almost equal in frequency, lock-in occurs at almost exactly that frequency.

### 3. Experimental beam-disc set-up and FE model

The mode lock-in test set-up consists of two subsystems, a rotating disc and a clamped beam on a slide-way (Fig. 1, left). The disc has a diameter of 358 mm and a thickness of 25 mm. The beam has a variable length from 0 mm to 200 mm to vary its natural frequencies and a square cross section of 10 mm. The normal load on the beam is generated by weight (adjustable from 0.5 kg to 5 kg) which presses the two subsystems together. The vibrations during squealing of the disc are measured by a triax laser vibrometer whereas those of the beam are recorded by

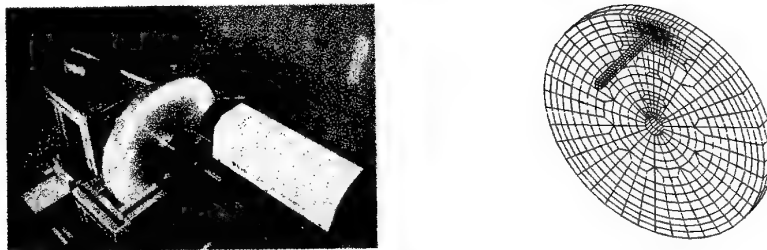


Figure 1: Experimental beam-disc set-up and FE model

a triax accelerometer. Noise generated by friction-induced vibrations is measured by a  $\frac{1}{4}$ "-microphone. All of the experiments are conducted with a direct aluminum to aluminum contact at the friction interface. The complete FE model, with its refined contact area, used to simulate mode lock-in, is shown in Fig. 1 (right). The domain of the beam-disc model is discretized by first-order hexahedron finite elements only. From the number of nodes (4582) and the additional degrees of freedom by contact, one obtains 14288 *DOF's* altogether. Rigid beams are fixed between the center and the inner diameter of the disc. By rotating the center nodes, a circular movement of the disc is generated. In order to bring both components of the friction couple into contact a force is applied at the end of the beam.

#### 4. FE model update

For improving the fit between simulated and experimental results, it is necessary to update the dynamic behaviour of the FE model corresponding to the experimental set-up. As a first step, both beam and disc are updated in free/free boundary conditions using natural frequencies and mode shapes measured by experimental modal analysis (EMA). Secondly, another EMA of the assembly is carried out. Using these data, the real acting boundary conditions and the modal damping values are taken into account by means of 3D springs and the Rayleigh damping approach in the numerical model. The model update is performed in a frequency range of 300 Hz to 20 kHz where besides the well-known out-of-plane modes several in-plane modes are identified as well. Fig. 2 shows the two mode shapes in assembled condition and identified by means of EMA which are involved when mode lock-in occurs.

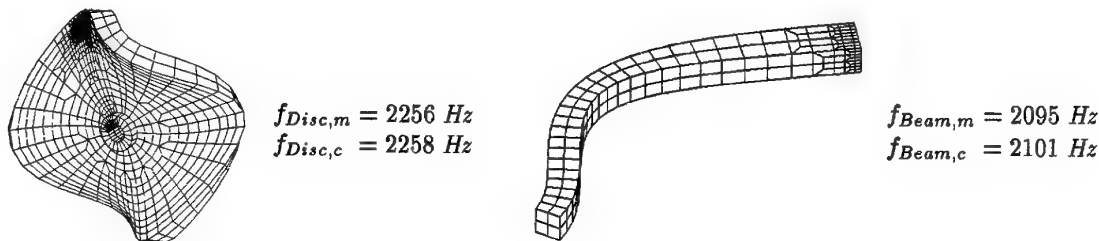


Figure 2: Squealing mode shapes of disc and beam (m=measured, c=calculated)

#### 5. Contact laws

In order to obtain realistic surface parameters for the simulation, a contact interface model [3] based on a statistical description of the surface roughness is used to derive a non-linear constitutive description for normal- and tangential contact. A number of assumptions are made: elastic contact of metallic surfaces; contact points do not interfere with each other; isotropic surface roughness; surface parameters do not change with time; dry friction.

**Normal contact:** Using these assumptions, each local contact area can be described by Hertzian theory. The normal force  $K_i$  to compress a summit  $i$  with normalized curvature  $s$  in contact with a plane is given by

$$K_i(\eta) = \frac{4}{3} E \frac{1}{\sqrt{\sigma_k s}} (\zeta - \eta)^{\frac{3}{2}} \sigma^{\frac{3}{2}} \quad (1)$$

with the normalized gap function  $\eta$ , Young's modulus  $E$ , root mean square value of the curvature  $\sigma_k$ , normalized height of the summit above a reference level  $\zeta$  and the root mean square value of the height distribution of summits  $\sigma$ .

**Tangential contact:** The tangential load-displacement function is formulated under the assumption of a total contact radius  $a$  calculated by Hertzian theory and a constant shear yield strength  $\tau_{max}$  of the junction. The tangential force  $Q_i$  of an asperity  $i$  with the normalized relative tangential displacement  $v$  is given by

$$Q_i(v, \eta) = \frac{\pi}{3} \tau_{max} a^2 \left[ (1 - \alpha^2) + 2 \frac{1 - \alpha^3}{\sqrt{1 - \alpha^2}} \right] \quad (2)$$

with the relative radius  $\alpha$  between slip and stick areas.

**Contact laws:** For the implementation in a macroscopic model the relations have to be distributed on the apparent area of contact. This is done via a statistical model of the surface, which assumes that the distribution of heights

is Gaussian. With the probability  $P$  for a summit with normalized curvature  $s$  and height  $\zeta$ , the apparent normal contact pressure  $p$  of all summits above  $\eta$  and the tangential contact stress  $\tau$  as a function of  $v$  and  $\eta$  are given by

$$p(\eta) = \frac{1}{h^2} \int_0^\infty \int_0^\infty K_i(\eta) P(\zeta, s) ds d\zeta, \quad \tau(v, \eta) = \frac{1}{h^2} \int_0^\infty \int_0^\infty Q_i(v, \eta) P(\zeta, s) ds d\zeta \quad (3)$$

with resolution  $h$  of a square grid of points for the surface roughness measurements.

Fig. 3 shows the constitutive laws of the contact interface using the equations above and surface roughness data of both beam and disc measured by laser scanning profilometry with a resolution  $h$  of  $1 \mu\text{m}$ . The material is aluminum and the material parameters are  $E = 70000 \frac{\text{N}}{\text{mm}^2}$ , Poisson's ratio  $\nu = 0.33$  and  $\tau_{max} = 385 \frac{\text{N}}{\text{mm}^2}$ .

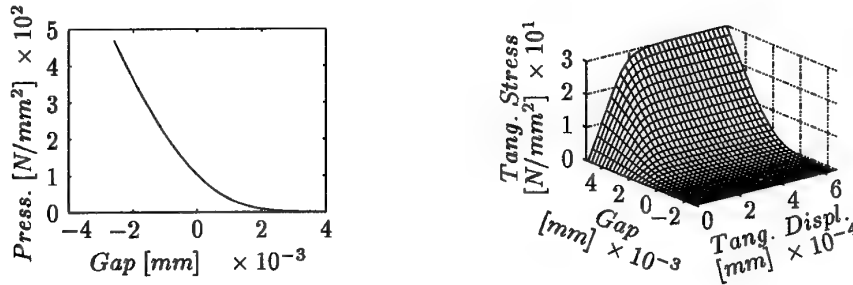


Figure 3: Normal and tangential contact law

The friction relation, as shown in Fig. 3 (right), is implemented in a standard FE code [2] by means of a user subroutine using a radial return mapping approach. Whereas the normal contact law (see Fig. 3, left) is defined as a piecewise-linear pressure-clearance relationship in tabular form.

## 6. Results

**Experimental results:** The experimental part is carried out using the test equipment shown in Fig. 1 (left). The test conditions are as follows: *beam length* = 151.5 mm (with a second beam eigenfrequency of 2095 Hz); *rotation speed* = 2.4 RPM (corresponding to an average disc velocity of  $40 \frac{\text{mm}}{\text{s}}$  at the contact location); *contact angle* =  $4^\circ$  (between normal vector to the disc and neutral axis of the beam); *normal load* = 15 N.

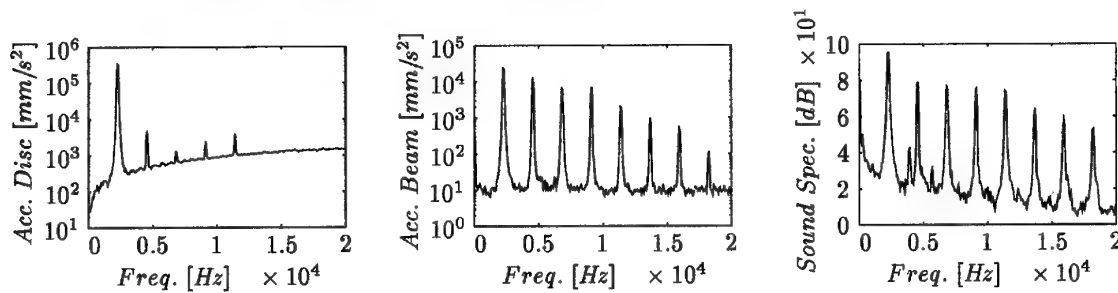


Figure 4: Spectrum from lock-in response for disc (left), beam (middle) and near field sound spectrum (right)

Fig. 4 shows an example of mode lock-in where the spectrum contains a dominant frequency and all of its higher harmonics. The lock-in frequency is determined by the third eigenfrequency of the disc (2256 Hz). However the experimental data show, that lock-in does not occur at exactly that frequency, but at a slightly lower one (2215 Hz). In contrast to the spectrum of the beam (Fig. 4, middle), which contains seven harmonics, the frequency response of the disc (Fig. 4, left) shows four harmonics only in the frequency range considered. Furthermore, it can be observed that the amplitudes of the beam harmonics decrease monotonically, whereas the amplitudes of the disc harmonics show a completely different behaviour. To verify the results described above, squealing generated during mode lock-in (Fig. 4, right) was measured. The sound spectrum has likewise a dominant frequency and higher harmonics.

**Numerical results:** In the following section the FE model developed previously is applied to a non-linear time-marching solution. Beam length, rotation speed, contact angle and normal load are equal to the values used in the experimental part whereas the friction relation, found in Section 5, is varied.

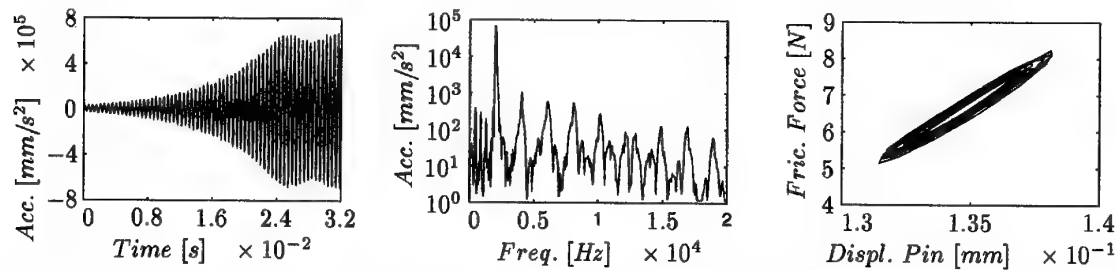


Figure 5: Onset of friction-induced vibrations in time domain (left), frequency domain (middle) and considerations on excitation energy (right)

Fig. 5 depicts the simulation results using the equivalently scaled friction relation ( $\mu(p = 14 \frac{N}{mm^2}) = 0.45$ ) if mode lock-in takes place. Both beam and disc vibrate at the same frequency, although the natural frequencies of the two sub-systems do not match ( $\Delta f = 157 \text{ Hz}$ ). It is evident that the oscillations of the system are actually friction-induced and self excited. They grow as a function of time (see Fig. 5, left) and reach a steady state after a certain period of time. By means of a FFT (Fig. 5, middle) it can be shown that the system responds with a dominant frequency (2017 Hz). The higher harmonics can be observed as well. Fig. 5 (right) shows the friction force versus relative displacement in lateral direction of the beam. The hysteresis area describes the amount of energy which is exciting the system during one single oscillation. It increases with increasing vibration amplitudes.

## 7. Conclusion

In this paper, an experimental and numerical study of the mode lock-in phenomenon, in particular of self-excited friction-induced oscillations was presented. A finite element model of the rotor-stator system, updated by means of experimental modal analysis was used to compute the dynamic behaviour of the experimental set-up. The numerical predictions for a mode lock-in case were compared with experimental results. A good correlation was found concerning the qualitative behaviour in frequency domain (dominant lock-in frequency and its higher harmonics). However the lock-in frequency numerically predicted was slightly lower than that experimentally observed. The unstable amplitude growth of the investigated beam-on-disc model configuration was underlined by calculating the excitation energy per vibration cycle.

Modelling friction-induced vibrations using the finite element method provides a simulation tool to obtain a deeper insight into the mechanisms of self-excited oscillations and guidance in avoiding them in the design of sliding systems. Although the analyzed model was comparably simple, the observations and methods used in this work can be applied to the analysis of more complex systems, for example disc brakes.

## Acknowledgements

The authors are indebted to DR.-ING. M. FISCHER and DR. RER. NAT. N. HOFFMANN at Robert Bosch GmbH (Corporate Research and Development, Research 2, Applied Physics-Acoustics) for being helpful and inspiring discussion partners.

## 8. References

- 1 AKAY, A., WICKERT, J., XU, Z.: Investigating criteria for the onset of mode lock-in, Pittsburgh, Carnegie Mellon University, Internal report, 1998.
- 2 HIBBIT, KARLSON, SORESEN: Abaqus Theory Manual 5.8, HKS, Inc., Pawtucket, 1998, 5.2.3.1–5.2.3.7.
- 3 WILLNER, K., GAUL, L.: A penalty approach for contact description by FEM based on interface physics, Contact Mechanics II (eds M. H. Aliabadi, C. Alessandri), Proceedings of the 2nd International Conference on Computational Methods in Contact Mechanics, CMP, Southampton, 1995, 257–264.

Addresses: R. ALLGAIER, Robert Bosch GmbH, Dept.: FV/FLP-Acoustics,  
Robert Bosch Platz 1, D-70839 Gerlingen-Schillerhöhe.  
PROF. L. GAUL, DR.-ING. K. WILLNER, University of Stuttgart, Institut A of Mechanics,  
Pfaffenwaldring 9, D-70550 Stuttgart.  
DR. RER. NAT. W. KEIPER, Bosch Braking Systems, Dept.: K1/NVH,  
126 rue de Stalingrad, F-93700 Drancy.

AMS A., SCHMALFUSS C., WEDIG W.

## Experimentelle und theoretische Untersuchungen an Scheibenbremsen

Die an einer Hydropulsanlage gewonnenen Ergebnisse über das Materialverhalten von Bremsbelägen werden im ersten Teil diskutiert. Der zweite Teil zeigt theoretische Untersuchungen an einem zeitvarianten Modell mit einem Freiheitsgrad. Der Tauselfehler der Scheibe und Oberflächenrauigkeiten werden als harmonische und stochastische Anregungen formuliert. Mit Hilfe von nichtlinearen Markov-Modellen können stochastische Anregungen mit begrenzten Amplituden und Frequenzen für Scheibenschlag, Reibwert und Oberfläche formuliert werden. Mittels des größten Lyapunov-Exponenten lassen sich Stabilitätsaussagen durchführen.

### 1. Experimentelle Untersuchungen an Bremsbelägen

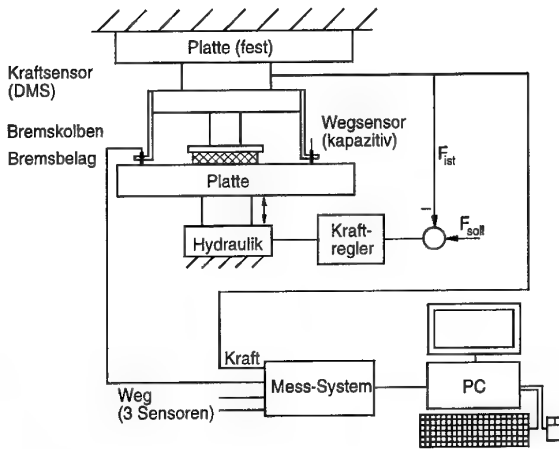


Bild 1: Hydropulser

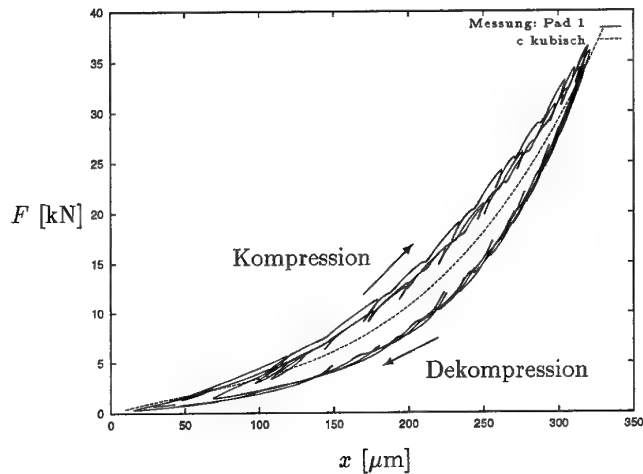


Bild 2: Steifigkeitskennlinie

Das Materialverhalten von Bremsbelägen wurde an einem Hydropulser (siehe Bild 1) experimentell ermittelt. Hierzu wurden die Bremsbeläge einachsig (kraftregelt) belastet und über Sensoren die Kraft  $F$  und der Weg  $x$  gemessen. In Bild 2 ist für eine quasi-statische Belastung eine Messung dargestellt. Gut zu erkennen ist der nichtlineare Verlauf der Steifigkeitskennlinie. Bei ansteigender Belastung (Kompression) wird die obere und bei abfallender Belastung (Dekompression) die untere Steifigkeitskennlinie durchlaufen. Mit einem Polynom dritten Grades

$$F(x) = a_1 x + a_2 x^2 + a_3 x^3 \quad (1)$$

kann die Steifigkeitskennlinie über ein Least-Square Verfahren approximiert werden. Für die Koeffizienten des Polynoms ergeben sich folgende Werte:

$$a_1 = 4.079 \cdot 10^{-2} \left[ \frac{\text{kN}}{\mu\text{m}} \right], \quad a_2 = -1.893 \cdot 10^{-4} \left[ \frac{\text{kN}}{(\mu\text{m})^2} \right], \quad a_3 = 1.261 \cdot 10^{-6} \left[ \frac{\text{kN}}{(\mu\text{m})^3} \right] \quad (2)$$

Wird als Anregung eine harmonische Kraft

$$F(t) = F_0 + F_1 \sin \omega t \quad (3)$$

mit dem Mittelwert  $F_0$ , der Amplitude  $F_1$  und der Erregerkreisfrequenz  $\omega$  verwendet, ergeben sich Kraft-Weg Verläufe, wie in Bild 3 dargestellt. Das visko-elastische Materialverhalten des Bremsbelages lässt sich durch das Modell

$$b \dot{x} + a_1 x + a_2 x^2 + a_3 x^3 = F(t) \quad (4)$$

mit dem konstanten Dämpfungsparameter  $b$  und der kubischen Steifigkeitskennlinie nach (1), beschreiben. Mit der Kraft  $F(t)$  und der Geschwindigkeit  $\dot{x}(t)$  kann die dissipative Arbeit  $\Delta E$

$$\Delta E = \int_0^{2\pi/\omega} F(t) dx = \int_0^{2\pi/\omega} F(t) \dot{x} dt \quad (5)$$

für eine Periode berechnet werden. Die stationäre Antwort des Systems (4) wird über eine numerische Integration berechnet. In Bild 4 ist der Quotient aus dissipativer Arbeit und dem Quadrat der Antwortamplitude  $\Delta E/A^2$  über der Frequenz  $f$  dargestellt. Für verschiedene Dämpfungswerte  $b$  sind die numerischen Ergebnisse eingezeichnet. Durch Vergleichen der Messung mit den Linien kann das Dämpfungsmaß  $b$  bestimmt werden. Im vorliegenden Fall liegt der Wert von  $b$  zwischen  $1 \cdot 10^{-5}$  und  $7 \cdot 10^{-5}$ .

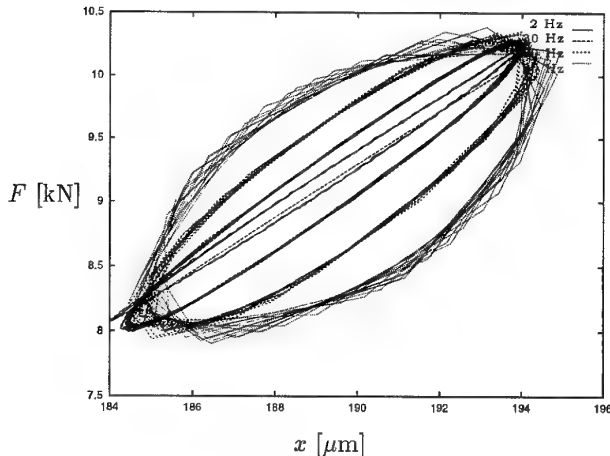


Bild 3: Harmonische Anregung

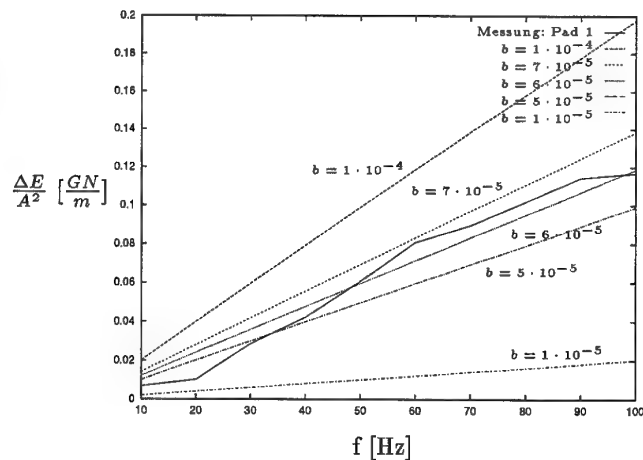


Bild 4: Dämpfungsparameter  $b$

## 2. Mechanisches Modell

Das in Bild 5 dargestellte mechanische Modell besteht aus einer starren Bremscheibe mit der Scheibenoberfläche  $u(s)$  und einem Bremsbelag mit der Masse  $m$ . Der Bremsbelag ist visko-elastisch (Federkonstante  $k$ , Dämpferkonstante  $d$ ) gelagert. Die Verbindung zwischen dem Belag und der Bremscheibe (1 Punktkontakt) wird durch ein visko-elastisches Element (Federkonstante  $c$ , Dämpferkonstante  $b$ ) hergestellt. Die Bewegung des Bremsbelages wird durch die Absolutkoordinate  $x$  beschrieben. Im Kontaktpunkt Scheibe/Belag herrscht Coulombsche Reibung (Reibwert  $\mu$ ), wobei die Richtung der Reibkraft stets der Tangente im Kontaktpunkt (Oberflächentangente) entspricht [1]. Die Bewegungsgleichung lautet:

$$\ddot{x} + 2D\omega_1\dot{x} + \omega_1^2 x + \frac{\mu + u'(s)}{1 - \mu u'(s)} [\omega_0^2 u(s) + 2B\omega_0 u'(s)\dot{s}] = 0 \quad (6)$$

mit der Scheibenkoordinate  $s$ , der Eigenkreisfrequenz  $\omega_1$ , dem Lehrschen Dämpfungsmaß  $D$ , den Parametern  $\omega_0$  und  $B$  sowie den Beziehungen

$$s = x - y, \quad \dot{s} = \dot{x} - \dot{y} \\ \omega_1^2 = k/m, \quad 2D\omega_1 = d/m, \quad \omega_0^2 = c/m, \quad 2B\omega_0 = b/m \quad (7)$$

## 3. Stochastischer Reibwert

Reibwertmessungen von Popp/Rudolph [2] zeigen deutlich eine Fluktuation des Reibwertes. Als mathematische Beschreibung für den Reibwert wird

$$\mu_t = \mu_0 + Z_t \quad (8)$$

mit konstantem Mittelwert  $\mu_0$  und additivem Rauschen  $Z_t$  angesetzt. Für  $Z_t$  kann das parametrische Modell

$$\dot{Z}_t = (\frac{1}{2}\sigma^2 - \omega_g)Z_t + \sigma\sqrt{Z_0^2 - Z_t^2}\dot{W}_t \quad (9)$$

mit dem Wiener-Prozeß  $W_t$ , der Rauschintensität  $\sigma$  und der Grenzfrequenz  $\omega_g$  verwendet werden. Das Modell besitzt die Eigenschaften:

$$\begin{aligned} |Z_t| &< z_0 && \text{Amplitudenbegrenzung} \\ R_z(\tau) &= \sigma_z^2 e^{-\omega_g |\tau|} && \text{Korrelation} \\ p(z) &= C(z_0^2 - z^2)^{(-1+\omega_g/\sigma^2)} \quad \text{für } \sigma^2 < \omega_g && \text{stationäre Verteilungsdichte} \end{aligned} \quad (10)$$

Über die Bedingung  $\int_{-z_0}^{+z_0} p(z) dz = 1$  kann die Konstante  $C$  bestimmt werden. Nach Transformation auf Polarkoordinaten mit der Amplitude  $A_t = \sqrt{X_t^2 + \dot{X}_t^2}$  und der Phase  $\Phi_t = \arctan \frac{\dot{X}_t}{X_t}$  ergeben sich für das homogene System (6) und für einen harmonischen Schlag  $u(t) = -u_0 - u_1 \sin \Omega v_0 t$  mit der statischen Vorspannung des Belages  $u_0$  und der Amplitude  $u_1$  folgende Ito-Gleichungen:

$$\begin{aligned} dA_t &= \{\sin \Phi_t \cos \Phi_t (1 - \omega_1^2) - \sin^2 \Phi_t (2D\omega_1 + 2B\omega_0 f_t)\} A_t dt \\ d\Phi_t &= \{-\sin^2 \Phi_t - \omega_1^2 \cos^2 \Phi_t - (2D\omega_1 + 2B\omega_0 f_t) \sin \Phi_t \cos \Phi_t\} dt \\ d\Psi_t &= \Omega v_0 dt \\ dZ_t &= -\omega_g Z_t dt + \sqrt{z_0^2 - Z_t^2} \sigma dW_t \\ \text{mit } f_t &= \frac{(\mu_0 + Z_t) u_1 \Omega \cos \Psi_t + u_1^2 \Omega^2 \cos^2 \Psi_t}{1 - (\mu_0 + Z_t) u_1 \Omega \cos \Psi_t} \end{aligned} \quad (11)$$

Mit dem multiplikativen Ergodentheorem (Oscelec) kann die asymptotische Stabilität der Ruhelage  $X = 0$  über

$$\begin{aligned} \lambda &= \lim_{t \rightarrow \infty} \frac{1}{t} \ln \frac{A_t}{A_0} = \lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t \{\sin \Phi_\tau \cos \Phi_\tau (1 - \omega_1^2) - \sin^2 \Phi_\tau (2D\omega_1 + 2B\omega_0 f_\tau)\} d\tau \\ \text{mit } f_\tau &= \frac{(\mu_0 + Z_\tau) u_1 \Omega \cos \Psi_\tau + u_1^2 \Omega^2 \cos^2 \Psi_\tau}{1 - (\mu_0 + Z_\tau) u_1 \Omega \cos \Psi_\tau} \end{aligned} \quad (12)$$

berechnet werden. Für die Parameter  $\omega_1 = 1$  [1/s],  $\omega_0 = 1000$  [1/s],  $B = 0.1$ ,  $D = 0.001$ ,  $\Omega = 5$  [1/m],  $\mu = 0.3$ ,  $z_0 = 0.1$ ,  $\omega_g = 10$  [1/s] und  $\sigma = 3$  ergibt sich die in Bild 6 gezeigte Stabilitätskarte. Für  $\lambda < 0$  ist das System stabil, für  $\lambda > 0$  instabil.

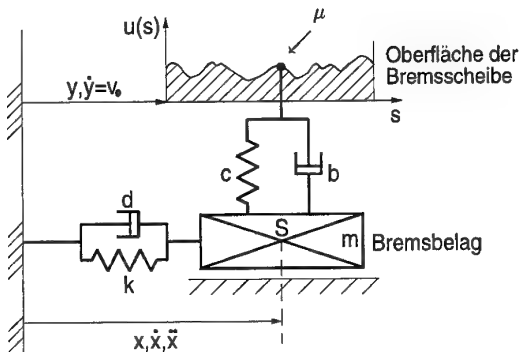


Bild 5: Mechanisches Modell

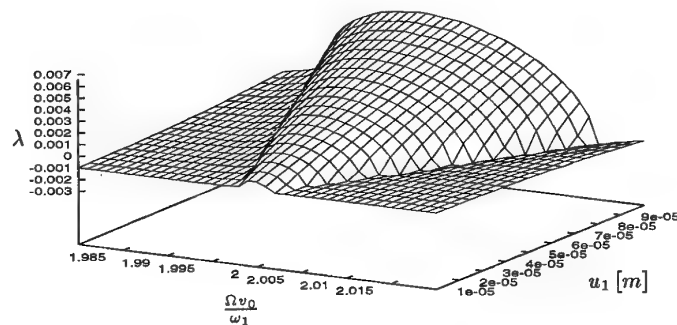


Bild 6: Stochastischer Reibwert (Lyapunov-Exponent  $\lambda$ )

#### 4. Stochastische Oberfläche

Die harmonische Anregung (Scheibenschlag) wird durch eine stochastische Anregung (Oberflächenrauigkeit) additiv überlagert. Für die Anregung ergibt sich somit

$$U_t' = u_1 \Omega \cos(\Omega v_0 t) + Z_t \quad (13)$$

mit dem parametrischen Modell  $Z_t$  nach (9). Nach Transformation auf Polarkoordinaten ergeben sich für das homo-

gene System (6) folgende Ito-Gleichungen:

$$\begin{aligned}
 dA_t &= \{\sin \Phi_t \cos \Phi_t (1 - \omega_1^2) - \sin^2 \Phi_t (2D\omega_1 + 2B\omega_0 f_t)\} A_t dt \\
 d\Phi_t &= \{-\sin^2 \Phi_t - \omega_1^2 \cos^2 \Phi_t - (2D\omega_1 + 2B\omega_0 f_t) \sin \Phi_t \cos \Phi_t\} dt \\
 d\Psi_t &= \Omega v_0 dt \\
 dZ_t &= -\omega_g Z_t dt + \sqrt{z_0^2 - Z_t^2} \sigma dW_t \\
 \text{mit } f_t &= \frac{\mu(u_1 \Omega \cos \Psi_t + Z_t) + (u_1 \Omega \cos \Psi_t + Z_t)^2}{1 - \mu(u_1 \Omega \cos \Psi_t + Z_t)}
 \end{aligned} \tag{14}$$

Über das multiplikative Ergodentheorem (Osceleddec)

$$\lambda = \lim_{t \rightarrow \infty} \frac{1}{t} \ln \frac{A_t}{A_0} \tag{15}$$

kann der größte Lyapunov-Exponent  $\lambda$  berechnet werden. In den Bildern 7 und 8 ist der größte Lyapunov-Exponent über der Amplitudenbegrenzung  $z_0$  aufgetragen. Verwendet wurden die Parameterwerte  $\omega_0 = 1000$  [1/s],  $\omega_1 = 1$  [1/s],  $\Omega = 5$  [1/m],  $\mu = 0.3$ ,  $B = 0.1$ ,  $D = 0.001$  und  $v_0 = 0.4$  [m/s]. In Bild 7 ist  $\omega_g = 10$  [1/s] und variiert wurde die Rauschintensität  $\sigma$  und der Scheibenschlag  $u_1$ . In Bild 8 zeigt sich der Einfluß der Korrelation (Grenzfrequenz  $\omega_g$ ) auf die Stabilität.

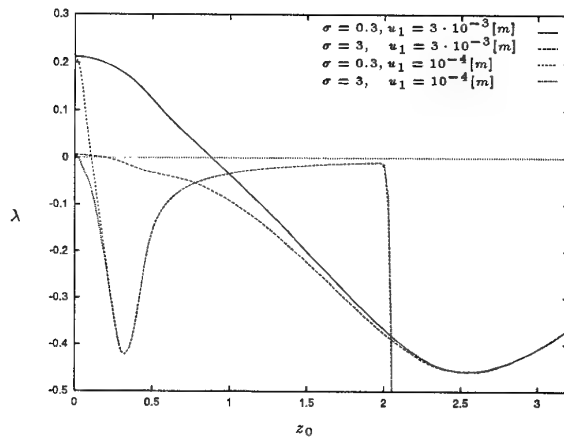


Bild 7: Einfluß von Schlag bzw. Rauschintensität

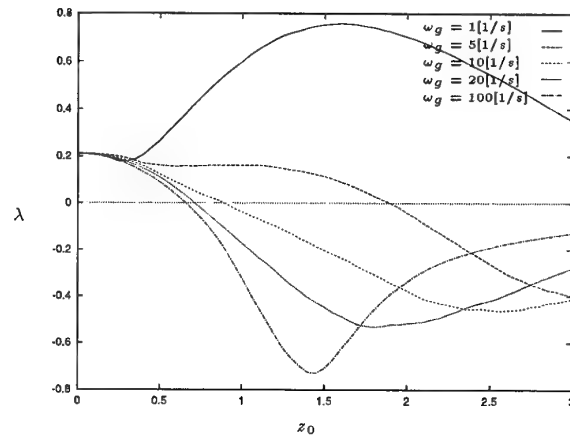


Bild 8: Einfluß der Korrelation

### Acknowledgements

Die Ergebnisse wurden durch das Bundesministerium für Bildung, Wissenschaft, Forschung und Technologie (BMBF) gefördert (FKZ 13N7176).

### 5. References

- 1 SCHMALFUSS, C.; WEDIG, W.; AMS, A.: Nichtlineare Oberflächenmodelle der Bremsscheibendynamik. ZAMM, Berlin 80 Suppl. 2, S. 299–300, ISSN 0946–8463
- 2 POPP, K.; RUDOLPH, M.: Beeinflussung und Vermeidung reibungsselbsterregter Schwingungen. BMBF-Zwischenbericht 2/98

Addresses: DR.-ING. A. AMS, DIPL.-MATH. TECHN. C. SCHMALFUSS, PROF. DR.-ING. W. WEDIG, Institut für Technische Mechanik, Universität Karlsruhe, Kaiserstr. 12, D-76128 Karlsruhe



ERTZ, M.; KNOTHE, K.

## Einfluss von Temperatur und Rauheit auf den Kraftschluss zwischen Rad und Schiene

*Gemessene Kraftschlusskennlinien zeigen charakteristische Abweichungen von Berechnungsergebnissen auf der Basis der Theorien von CARTER und KALKER. Die Unterschiede bestehen einerseits in einer geringeren Anfangssteigung, andererseits im Auftreten eines Maximums, nach dessen Überschreiten es zu einem teilweise erheblichen Abfall des Kraftschlusses kommt. In der vorliegenden Arbeit wird gezeigt, dass die Berücksichtigung von Mikrorauheiten zu einer realistischeren Anfangssteigung der berechneten Kennlinien führt. Weiterhin wird die Temperaturerhöhung aufgrund der in der Kontaktfläche auftretenden Reibleistung berechnet. Unter der Annahme einer temperaturabhängigen Reibungszahl kann damit auch der Abfall der Kennlinie bei hohen Schlüpfen erklärt werden.*

### 1. Kraftschlussberechnung nach CARTER

Der Kontakt zwischen Rad und Schiene ist ein dreidimensionales Problem. Unter der Wirkung einer Normalkraft  $N$  bildet sich eine Kontaktfläche aus, die in vielen Fällen nach der Theorie von HERTZ berechnet werden kann. Man erhält eine Ellipse mit den Halbachsen  $a$  (in Fahrtrichtung) und  $b$ . Für die Untersuchung des Spannungs- und Verzerrungszustandes unter einer Tangentialkraft  $T$  beim Antreiben oder Bremsen gehen wir auf den Kontakt eines Zylinders gegen eine Ebene über, womit sich ein zweidimensionales Modell ergibt. Die Größen  $a$  und  $p_{z0}$  (Maximalwert der Normalspannung) werden aus dem dreidimensionalen Normalkontakt übernommen. Dazu kommt hier noch die Breite  $L$  des Ersatzzylinders. Mit dieser Vereinfachung ist die Lösung des Tangentialkontaktproblems für glatte Oberflächen nach CARTER analytisch möglich. Es zeigt sich, dass eine Tangentialkraft immer mit dem Auftreten einer Relativgeschwindigkeit  $v_s = v_R - v_0$  zwischen der Fahrgeschwindigkeit  $v_0$  und der Umfangsgeschwindigkeit des Rades,  $v_R = \Omega_0 r$ , verbunden ist. Bezogen auf die mittlere Geschwindigkeit  $(v_0 + v_R)/2$  erhält man daraus den Schlupf  $\nu_x$ . Die Beziehung zwischen Kraftschluss  $f = T/N$  und Schlupf lautet nach CARTER

$$f(\nu_x) = \mu \left[ 2 \left( \frac{\nu_x}{\nu_{x,max}} \right) - \left( \frac{\nu_x}{\nu_{x,max}} \right)^2 \right], \quad (1)$$

solange  $\nu_x < \nu_{x,max}$  mit  $\nu_{x,max} = \mu a/r$  gilt. Dabei tritt am Auslaufrand der Kontaktfläche immer ein Gleitgebiet auf, dessen Ausdehnung mit steigendem Schlupf zunimmt, bis es für  $\nu_x = \nu_{x,max}$  die gesamte Kontaktfläche bedeckt. Damit ist gleichzeitig der maximale Kraftschluss  $f(\nu_{x,max}) = \mu$  erreicht, der sich bei weiterer Steigerung des Schlupfes nicht mehr ändert. Eine übersichtliche Darstellung dieser Zusammenhänge findet sich z. B. bei JOHNSON [4].

### 2. Berücksichtigung von Mikrorauheiten

Abbildung 1 zeigt einen Ausschnitt aus der Profilmessung auf einer Schienenlaufläche. Die Lösung von Normal- und Tangentialkontaktproblem ist bei solchen Oberflächen nur numerisch möglich. In einer Arbeit von KNOTHE und THEILER [5], deren Ergebnisse hier wiedergegeben werden, wird dazu ein spezielles Randelement-Verfahren verwendet. Für diese Berechnungen werden die Rauheiten von Rad und Schiene zusammengefasst und einem der beiden Körper zugeschlagen, während der andere als glatt angenommen wird. Bei der Lösung des Normalkontaktproblems zeigt sich, dass die Berührung der Oberflächen nur in den Spitzen der Rauheiten stattfindet, wo die maximalen Druckspannungen ein Vielfaches der Werte beim Kontakt glatter Oberflächen erreichen können. Als Maß für die Auswirkung der Rauheiten wird das Verhältnis  $A_{rau}/A_{glatt}$  herangezogen.

Zur Lösung des Tangentialkontaktproblems ist eine weitere Vereinfachung erforderlich. Es wird angenommen, dass die Verschiebungen im Gleitbereich so klein sind, dass der Kontaktvorgang als stationär betrachtet werden kann. Eine Änderung der Kontaktfläche durch das Gleiten der Rauheitsspitzen wird ebenso vernachlässigt wie die Bewegung der gesamten Oberflächen mit der Fahrgeschwindigkeit. In Abbildung 1 sind die numerisch berechneten Kraftschlusskennlinien für unterschiedliche Rauheiten dargestellt. Es zeigt sich, dass die Anfangssteigung mit zunehmender Rauheit abnimmt und das durch die Reibungszahl gegebene Maximum entsprechend später erreicht

wird. Mit guter Näherung kann hierfür

$$f_{rau}(\nu_x) = f_{glatt} \left( \frac{A_{rau}}{A_{glatt}} \nu_x \right) \quad (2)$$

geschrieben werden. Wegen der Vernachlässigung des Gleitens von Rauheitsspitzen ist es nicht sinnvoll, die auf diese Weise berechneten Kraftschlusskennlinien für raue Oberflächen in den Bereich vollen Gleitens fortzusetzen.

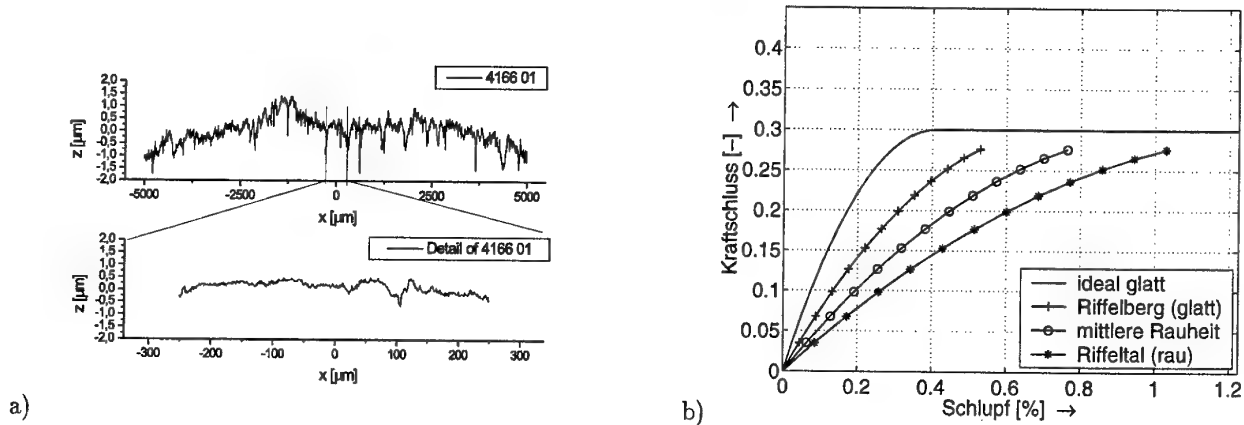


Abbildung 1: Ausschnitt aus der Profilmessung auf einer Schienenlaufläche (a) und numerisch berechnete Kraftschlusskennlinien für unterschiedliche Rauheiten (b)

### 3. Temperaturentwicklung im Radsatz

Wenn ein angetriebenes, auf der Schiene rollendes Rad die Tangentialkraft  $T = fN$  überträgt und dabei die Gleitgeschwindigkeit  $v_s = v_R - v_0$  auftritt, dann ergibt sich daraus die Reibleistung  $P_{reib} = fNv_s$ . Wir nehmen an, dass sie unmittelbar in der Kontaktfläche vollständig in den Wärmestrom  $\dot{Q}_{reib} = P_{reib}$  umgewandelt wird, der in das Innere von Rad und Schiene abfließt. Davon geht der Anteil  $\dot{Q}_{reib,R} = \delta \dot{Q}_{reib}$  mit

$$\delta = \frac{\beta_R \sqrt{v_R}}{\beta_R \sqrt{v_R} + \beta_S \sqrt{v_0}} \quad (3)$$

in das Rad.  $\beta = \sqrt{\lambda \rho c}$  ist der materialabhängige Wärmeeindringkoeffizient. Die Aufteilung des Wärmestroms ergibt sich aus der Forderung, dass an der Oberfläche von Rad und Schiene im Kontakt Temperaturgleichheit herrschen muss. Aufgrund der kurzen Kontaktzeit tritt Wärmeleitung nur senkrecht zur Kontaktfläche auf. In Abbildung 2 sind die Oberflächentemperaturen dargestellt, die sich bei gleicher Reibleistung  $\dot{Q}_{reib}$  in Abhängigkeit von der lokalen Verteilung der Reibleistungsdichte  $\dot{q}_{reib}(x) = \mu(x)p_z(x)v_s(x)$  ergeben. Der Einfluss beschränkt sich im Wesentlichen auf den Ort der Maximaltemperatur, deren Wert aber kaum von der Verteilung abhängt. Ausserhalb des Kontaktes gleichen sich die Temperaturverläufe sehr schnell wieder an. Mit jeder Umdrehung wird dem Rad die Wärmemenge

$$q_{reib,R} = \frac{\delta}{Lv_R} \dot{Q}_{reib} \quad (4)$$

pro Flächeneinheit zugeführt. Nach mehreren Umdrehungen des Rades steigt seine Temperatur durch diese regelmäßige Wärmezufuhr immer weiter an, während die Schiene weiterhin mit Umgebungstemperatur in den Kontakt eintritt. Damit ergibt sich ein Wärmestrom aus dem warmen Rad in die kalte Schiene, der mit guter Näherung als Temperaturengleich von zwei halbinendlichen Körpern mit verschiedenen, jeweils konstanten Anfangstemperaturen berechnet werden kann. Mit der Eintrittstemperatur  $\Theta_{R0}$  des Rades stellt sich im Kontakt spontan eine zeitlich nicht veränderliche Kontakttemperatur  $\Theta_m = \delta \Theta_{R0}$  mit  $\delta$  wie in (3) ein. Die gesamte durch Temperaturdifferenz während einer Umdrehung übertragene Wärmemenge pro Flächeneinheit ist damit

$$q_{temp,R} = -\frac{\delta \beta_S}{v_R} \sqrt{\frac{8av_0}{\pi}} \Theta_{R0}. \quad (5)$$

Wie sich mit einer einfachen Abschätzung zeigen lässt, ist die Wärmeübertragung durch Konvektion an die Umgebungsluft im Vergleich mit der Wärmeleitung in die Schiene von untergeordneter Bedeutung.

Damit sich die Temperatur des Rades bis auf periodische Schwankungen während einer Umdrehung nicht mehr ändert, muss der Wärmezufuss aus Reibleistung gleich dem Wärmeabfluss aus Temperaturdifferenz sein. Aus (4) und (5) ergibt sich damit die Temperatur, in der thermisches Gleichgewicht vorliegt, zu

$$\Theta_{\infty} = \frac{1}{\beta_S L} \sqrt{\frac{\pi}{8av_0}} \dot{Q}_{reib}. \quad (6)$$

In Abbildung 2 erkennt man, dass die Schwankungen um diesen Wert auf die unmittelbare Umgebung des Kontaktes beschränkt sind. Die Gleichgewichtstemperatur ist etwa doppelt so hoch wie die mittlere Kontakttemperatur beim einmaligen Kontaktdurchlauf.

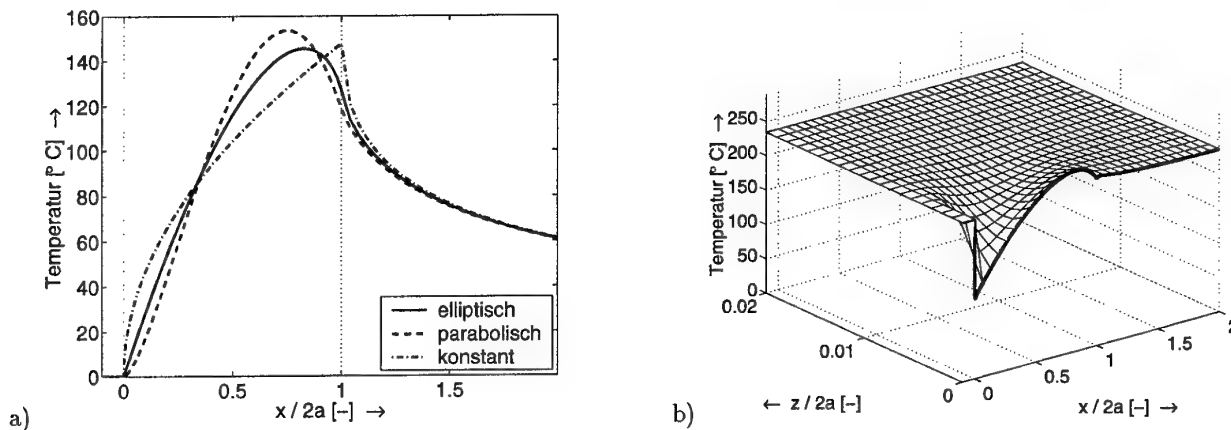


Abbildung 2: Oberflächentemperatur beim einmaligen Kontaktdurchlauf für unterschiedliche Verläufe der Reibleistungsdichte (a) und Temperaturverlauf im thermischen Gleichgewicht (b)

#### 4. Temperaturabhängigkeit der Reibungszahl

Aus den vorangegangenen Ausführungen folgt, dass es in der Kontaktfläche zum Gleiten zwischen metallischen Oberflächen kommt und dabei gleichzeitig erhebliche Temperaturerhöhungen auftreten. Als mögliche Auswirkung der Temperaturerhöhung wird im Folgenden eine mit zunehmender Temperatur abfallende Reibungszahl untersucht. Dafür lassen sich unterschiedliche physikalische Erklärungen angeben. BOWDEN und TABOR [1] nehmen an, dass beim Gleitvorgang ein ständiges Verschweißen und Aufbrechen von Metallbrücken stattfindet. Der Bewegungswiderstand, d. h. die Reibungszahl, wird damit durch die Scherfestigkeit der entstandenen Verbindungen bestimmt und die Temperaturabhängigkeit der Reibungszahl ergibt sich aus der Temperaturabhängigkeit der Scherfestigkeit. Bei KRAFT [6] findet sich eine andere Begründung, die auf einem atomaren Modell beruht. Wir gehen hier nach BOWDEN und TABOR von der Streckgrenze eines mittelfesten Stahls aus, die einen monotonen Abfall mit zunehmender Temperatur aufweist [3]. Für nicht allzu hohe Temperaturen kann mit guter Näherung linearisiert werden. Mit der Übertragung der Temperaturabhängigkeit auf die Reibungszahl lässt sich dann die Beziehung

$$\mu(\Theta) = \mu_0 \left( 1 - \frac{\Theta}{\Theta_{\mu}} \right) \quad (7)$$

mit  $\mu_0$  als Reibungszahl bei 0° C angeben. Der Parameter  $\Theta_{\mu}$  ist ein Maß für die Steigung der Kennlinie und wird für die folgenden Rechnungen mit 500° C angenommen. Die Linearisierung ist bis etwa 250° C zulässig.

Für die folgenden Untersuchungen wird angenommen, dass die Reibungszahl über die gesamte Kontaktfläche konstant ist und nur von der mittleren Kontakttemperatur abhängt.

#### 5. Temperaturabhängige Kraftschlusskennlinie

Bei vorgegebenem Schlupf kommt es zu einem Zusammenwirken von Kraftschluss, Reibleistung, Temperaturerhöhung und temperaturabhängiger Reibungszahl. Wir betrachten den Fall vollen Gleitens, wobei der Kraftschluss gleich der Reibungszahl ist. Ausgangspunkt ist die temperaturabhängige Reibungszahl (7). Wenn man hier den Zusammenhang zwischen der mittleren Temperatur und der Reibleistung (6) einsetzt und die Reibleistung  $\dot{Q}_{reib} = f N v_s$  wiederum

auf den Kraftschluss zurückführt, erhält man die Gleichung

$$f(v_s) = \frac{\mu_0}{1 + v_s/v_\infty} \quad (8)$$

für eine mit zunehmender Gleitgeschwindigkeit abfallende Kraftschlusskennlinie. Die hierbei verwendete Abkürzung

$$v_\infty = \frac{\Theta_\mu \beta_S L}{\mu_0 N} \sqrt{\frac{8av_0}{\pi}} \quad (9)$$

besitzt die Dimension einer Geschwindigkeit. In gleicher Weise lässt sich auch die Cartersche Lösung (1) mit einer von der mittleren Kontakttemperatur abhängigen Reibungszahl berechnen. Die Ergebnisse sind in Abbildung 3 dargestellt. Diese Kennlinie beschreibt den Kraftschluss, der sich nach hinreichend langer Zeit einstellt, wenn thermisches Gleichgewicht erreicht worden ist. Die Vorgehensweise lässt sich zumindest qualitativ auch auf Übergangszustände anwenden. Da die Temperatur für gegebene Betriebsbedingungen ihr Maximum im thermischen Gleichgewicht erreicht, ist der Temperatureinfluss und damit der Abfall der Kennlinie in diesen Fällen geringer.

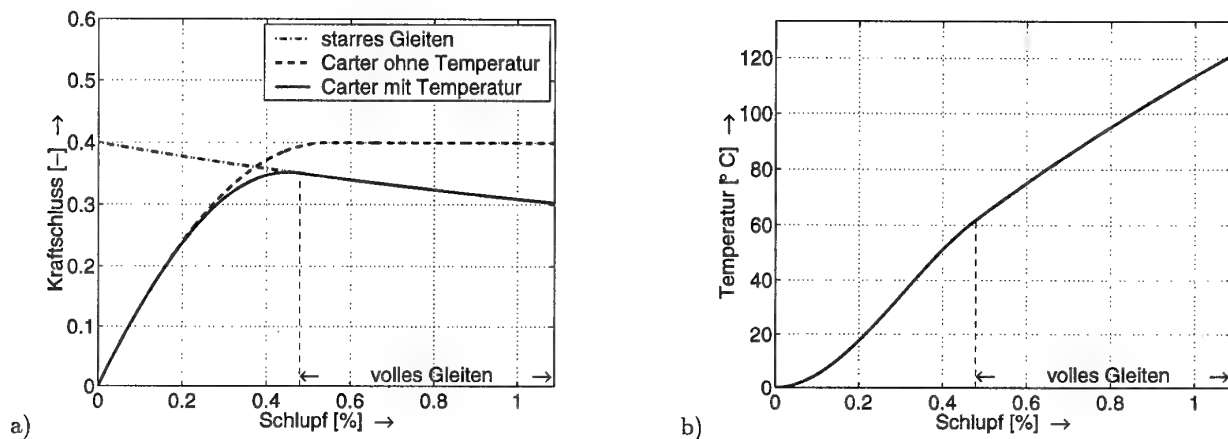


Abbildung 3: Kraftschlusskennlinie bei temperaturabhängiger Reibungszahl (a) und Temperatur im thermischen Gleichgewicht (b), Fahrgeschwindigkeit  $v_0 = 75$  m/s

## 6. Zusammenfassung

Die vorgestellten Ergebnisse liefern physikalisch begründete Erklärungen für die charakteristischen Abweichungen zwischen gemessenen und berechneten Kraftschlusskennlinien. Mit der Berücksichtigung von Mikrorauheiten nimmt die Anfangssteigung der berechneten Kennlinien ab. Durch die Berechnung der Kontakttemperatur in Verbindung mit der Annahme einer temperaturabhängigen Reibungszahl erhält man einen ausgeprägten Abfall der Kennlinie bei hohen Schlüpfen. Beide Effekte führen zu einer besseren Übereinstimmung mit gemessenen Kurven [2].

Diese Arbeit wurde im Rahmen des Sfb 605 „Elementarreibereignisse“ an der TU Berlin durchgeführt.

## 7. Literatur

- 1 BOWDEN, F. P.; TABOR, D.: Reibung und Schmierung fester Körper. Springer, Berlin, Göttingen, Heidelberg, 1959.
- 2 HÄSE, P.; MENTH, S.: Kraftschluss bei Triebfahrzeugen - Modellbildung und Verifikation an Messdaten. In: Elektrische Bahnen 94 (1996), Nr. 5, 125-134.
- 3 Handbuch der Kennwerte von metallischen Werkstoffen. Deutscher Verlag für Schweißtechnik, Düsseldorf, 1997.
- 4 JOHNSON, K. L.: Contact Mechanics. Cambridge University Press, Cambridge, 1985.
- 5 KNOTHE, K.; THEILER, A.: Normal and tangential contact problem with rough surfaces. In: I. Zobory (Ed.), Proc. of the 2nd Mini Conference on Contact Mechanics and Wear of Rail/Wheel Systems held at Budapest in August 1996, 34-43, Budapest, 1997.
- 6 KRAFT, K.: Der Einfluß der Fahrgeschwindigkeit auf den Haftwert zwischen Rad und Schiene. Archiv für Eisenbahntechnik, Folge 22 (1968), 58-78.

Adresse: DIPL.-ING. MARTIN ERTZ, PROF. DR.-ING. KLAUS KNOTHE, Technische Universität Berlin, Institut für Luft- und Raumfahrt, Marchstr. 12, 10587 Berlin

MÜLLER, S. AND R. KÖGEL

## Numerical simulation of roll-slip oscillations in locomotive drives

*This paper deals with roll-slip oscillations in locomotive drives. Numerical simulation results are presented which demonstrate that roll-slip oscillations may arise if the contact conditions at the left and right wheel of a wheelset are different. Even if an adhesion controller can detect and reduce these vibrations within a few seconds the drive chain is still subject to high loads, there are noisy vibrations in the locomotive car body and the rail surface is damaged by periodic wear patterns. Usually, it is tried to avoid roll-slip oscillations by an optimal adhesion controller design. A different approach is presented in this paper. Instead of optimising a controller it is demonstrated how mechanical design parameters influence the proneness to roll-slip oscillations and for one locomotive type it is shown how a certain mechanical parameter combination helps to avoid roll-slip.*

### 1. Introduction

Roll-slip oscillation in locomotive drives is an unstable vehicle motion and is usually associated with a slip velocity-traction force characteristic with a negative gradient at relatively high slip velocities. This characteristic was found during many measurements [1] [2]. Operating the locomotive where the gradient of the slip velocity-traction force curve is negative introduces negative damping to the system and the vehicle dynamics can become unstable. This causes high mechanical loads in the drive chain, noisy vibrations in the locomotive car body and periodic wear patterns on the rail surface.

Presumably, in [3] roll-slip oscillations in locomotive drives have been first thoroughly investigated using a simple numerical model of the mechanical drive chain. More recently, [4] and [5] are dealing with roll-slip oscillation and how control concepts can help to avoid this phenomenon. In this paper a different approach is presented. In addition to an already existing control strategy [6] which detects and reduces roll-slip oscillations it is tried to reduce the proneness to roll-slip by the mechanical design of the locomotive drive.

For one locomotive type numerical simulation results demonstrate that roll-slip oscillations may arise if the contact conditions at the left and right wheel of a wheelset are different. It is also shown that the adhesion controller of the locomotive detects these unstable vibrations and the controller reduces the high amplitudes after a few seconds. But during this period of time the drive chain is subject to high mechanical loads. A mechanical design which had a smaller proneness to roll-slip would reduce the number of roll-slip occurrences during operation. This would increase the lifetime of the drive components, reduces the noise inconveniences and preserves the track. It is therefore investigated how mechanical design parameters influence the proneness to roll-slip and for the locomotive type investigated it is shown how a certain parameter combination helps to avoid roll-slip.

### 2. The numerical simulation model

The modeling of the control part of the locomotive comprises the motor control and the adhesion control. The motor control is considered by measured transfer functions. The adhesion controller controls the slip velocity, which is the relative velocity between wheel and rail in the point of contact. It determines the optimal slip where maximal traction forces are obtained [6]: A speed sensor measures the response of the angular velocity of the rotor to the motor torque plus a sinusoidal test signal added to the motor torque. Based on the phase between filtered sinusoidal rotor speed response and the sinusoidal motor torque test signal the gradient of the slip velocity-traction force curve at the operating point of the drive is computed. The actual gradient is compared to a predefined desired gradient. If the difference between both is too large a new desired rotational rotor speed is calculated which changes the motor torque.

The mechanical structure has been modeled using a Multi-Body-System program. All model parts are rigid and all springs and dampers are linear. Under consideration of structural symmetry it comprises half of the car body, one bogie, secondary and primary support, drive components and three wheelsets. The elasticity of the wheelset axle about the lateral axis is considered by two torsional springs. One between gear wheel and right and left wheel, respectively. The locomotive type investigated is driven by a so-called nose-suspended drive, which consists of motor

box, rotor, gear wheel and the motor box support. The gearwheel is fixed to the wheelset axle and is in the vicinity of the right wheel. The motor box is supported by the wheelset axle but can rotate about it. The motor box is elastically connected to the bogie by the motor box support.

For the numerical simulation a mechatronical model has been set up within MATLAB/SIMULINK which comprises the mechanical structure, the control part and the interaction between mechanics and control.

### 3. Simulation results if contact conditions at left and right wheel are different

With the mechatronical model a time step integration has been performed to simulate the dynamical behaviour of a locomotive if the contact conditions at left and right wheel of a wheelset are different. In Figure 1 the slip velocity-traction force characteristics are shown which have been assumed for the numerical simulation.

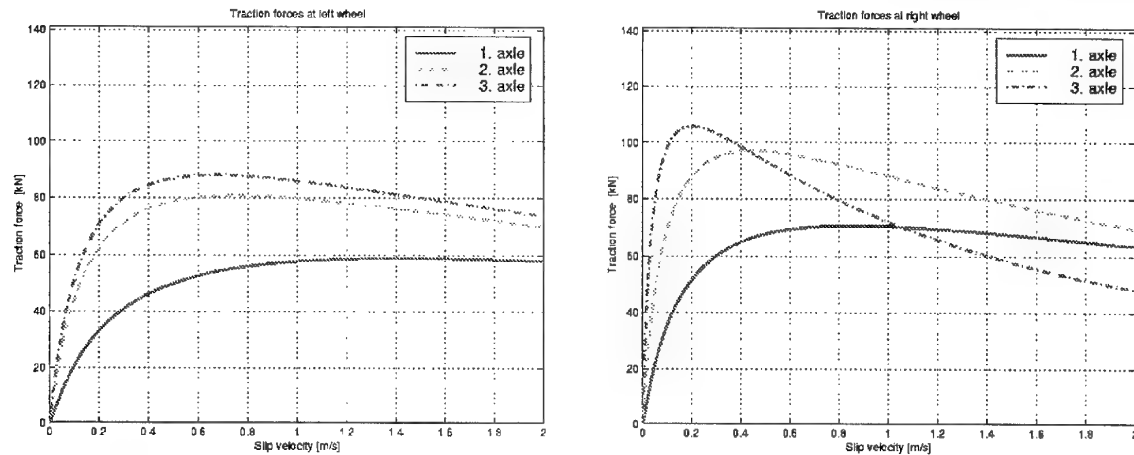


Figure 1: (a) Longitudinal contact condition at left and right wheel if  $t < 10$  sec, (b) longitudinal contact condition at right wheel after ten seconds

At the beginning of the simulation the contact is the same for both wheels of a wheelset. But the contact is different at different wheelsets as soon as conditioning effects caused by a preceding wheelset are considered (Figure 1(a)). After ten seconds the contact at the right wheel changes (Figure 1(b)). This may happen if one side of the track is frosted or covered by oil or dirt. Contact conditions can also be different due to kinematic effects: In the curve the rolling radii of left and right wheel are different and the translational velocities of inner and outer wheel are not the same. This results in different slip velocities at the right and left wheel of a wheelset and thus in different operating points on the slip velocity-traction force curve.

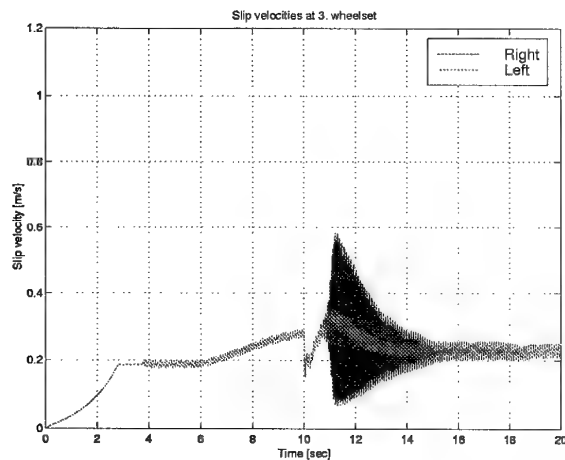


Figure 2: Slip velocity at left and right wheel of the third wheelset

In Figure 2 the outcome of the numerical simulation for the slip velocity at the left and right wheel of the third

wheelset is plotted. At the beginning the locomotive starts and the slip velocity increases till a predefined limit of 0.2 m/s is obtained. After four seconds the adhesion controller superposes a test signal of 11 Hz to the motor torque. The adhesion controller starts to change the slip velocity after 6 seconds until the right contact changes at  $t = 10$  sec. This causes a sudden drop which is followed by a new increase until the vehicle dynamics become unstable and the slip velocity at the right wheel is oscillating with high amplitudes and a frequency of about 147 Hz. An eigenvalue analysis of the mechanical structure reveals that at this frequency the right wheel is rotating about the lateral axis. The rotation of right wheel and gearwheel is out-of-phase, the amplitudes of the rotation of the left wheel are small. The roll-slip oscillation arises since the right wheel of the third wheelset is operating at slip-velocities where the gradient of the slip velocity-traction force curve is negative. The adhesion control detects the unstable motion and reduces the slip velocity until the gradient of the slip velocity-traction force curve is positive at all wheels. The amplitudes thus decrease and after a few seconds the vehicle dynamics again is stable.

#### 4. Investigation of the influence of mechanical design parameters

In Figure 2 it is shown that the vehicle dynamics may become unstable if the contact conditions at the left and right wheel of a wheelset are different. It is also demonstrated that the adhesion controller detects the unstable motion and reduces the amplitudes of the oscillations after a few seconds. Nevertheless, the drive chain is subject to high mechanical loads during this period of time. It is now investigated how mechanical design parameters influence the proneness to roll-slip and for the locomotive type investigated mechanical design parameters are given which are likely to reduce the number of roll-slip occurrences.

For the investigation of the proneness to roll-slip the torsional stiffness and damping of the connection gearwheel-wheelset axle and the torsional stiffness of the wheelset axle between gearwheel and right wheel of a wheelset are varied. The investigation is performed in the frequency domain. Except for the right wheel of the third wheelset the gradient of the slip velocity-coefficient of friction curve is zero. At the right wheel of the third wheelset the gradient is varied from zero to a negative value. Note, that for the traction force at one wheel the relation  $F_{traction} = \mu(\Delta v) N_{static}$  is assumed, where  $\mu$  is the coefficient of friction,  $\Delta v$  is the slip velocity and  $N_{static}$  is the static wheel load in vertical direction.

For each mechanical design parameter combination the eigenvalues of the linearized mechatronical model are calculated for different gradients at the right wheel of the third wheelset and the threshold gradient is determined, which is the minimal gradient where the vehicle dynamics is still stable. The outcome of this investigation is summarized in Figure 3(a), Figure 3(b) and Figure 4.

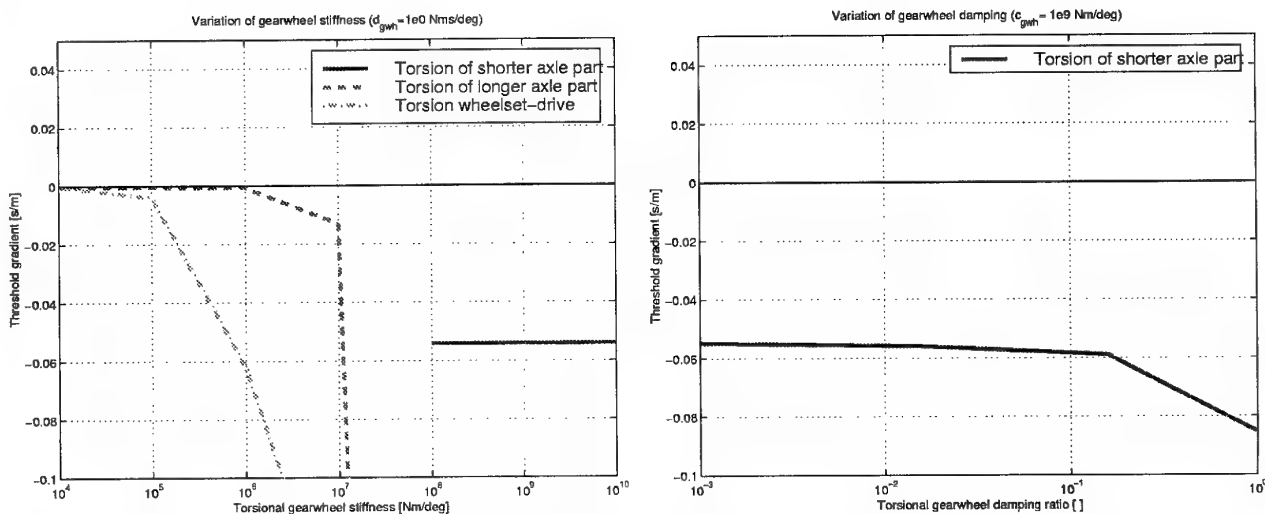


Figure 3: Threshold gradient at (a) different torsional stiffnesses and (b) different torsional damping coefficients of the connection gearwheel-wheelset axle

In Figure 3(a) the threshold gradient is plotted at different torsional stiffnesses between gearwheel and wheelset axle. At relatively small stiffnesses the proneness to roll-slip is high. Even for gradients only slightly smaller than zero the system becomes unstable. The corresponding natural mode is characterized by a torsion of the wheelset axle

between gearwheel and left wheel. At higher stiffnesses the proneness to roll-slip clearly decreases and the vehicle dynamics becomes unstable if the gradient at the right wheel of the third wheelset is smaller than  $-0.055$  s/m. The unstable motion is then characterized by a torsion of the wheelset axle between gearwheel and right wheel.

The threshold gradient at different torsional damping coefficients of the connection gearwheel-wheelset axle is plotted in Figure 3(b). For this calculation the torsional stiffness between gearwheel and wheelset axle has been set to  $1 \times 10^9$  Nm/deg. Along the x-axis a damping ratio is plotted which has been defined as  $D_{gwh} = 0.5 d_{gwh} / \sqrt{c_{gwh}}$ , where  $d_{gwh}$  is the torsional damping coefficient and  $c_{gwh}$  is the torsional stiffness between gearwheel and wheelset axle. It reveals that the proneness to roll-slip decreases if the torsional damping increases.

In Figure 4 the threshold gradient is plotted at different torsional stiffnesses of the wheelset axle between gearwheel and right wheel. For this calculation the torsional stiffness of the connection gearwheel-wheelset axle is  $1 \times 10^9$  Nm/deg, the corresponding damping coefficient is  $1 \times 10^0$  Nms/deg. It is illustrated that the vehicle dynamics is most likely to remain stable if the torsional stiffness is comparatively high.

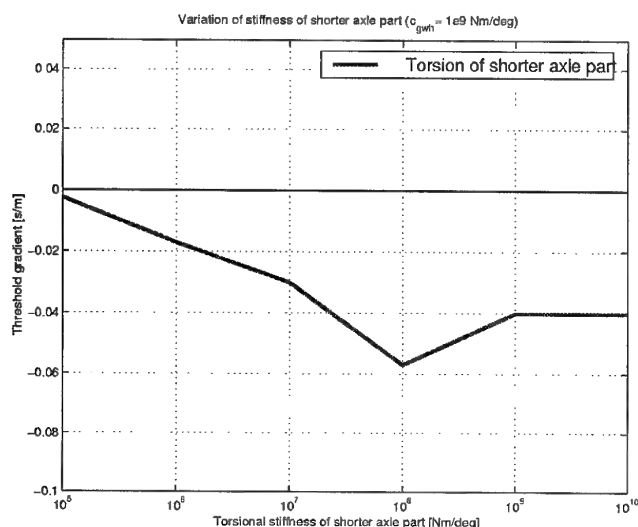


Figure 4: Threshold gradient at different torsional stiffnesses of the wheelset axle between gearwheel and right wheel

The outcome of the variation of mechanical design parameters is that the locomotive type investigated has a reduced proneness to roll-slip if the connection between gearwheel and wheelset axle is stiff and damped and if the torsional stiffness of the wheelset axle is relatively high. With these mechanical design parameters it is therefore likely to reduce the number of roll-slip occurrences during operation.

### Acknowledgements

The authors wish to express their appreciation to DaimlerChrysler Rail Systems, Switzerland which financed the development of the numerical simulation model used for the investigations presented in this paper.

### 5. References

- LOGSTON, C.F., JR. AND G.S. ITAMI: Locomotive Friction-Creep Studies, ASME Paper No. 80-RT-1, April, 1980.
- HÄSE, P. AND S. MENTH: Kraftschluss bei Triebfahrzeugen - Modellbildung und Verifikation an Messdaten, Elektrische Bahnen 94, 1996.
- E. KÖRNER: Reibschwingungen eines elektrischen Triebfahrzeuges an der Haftwertgrenze, ZEV-Glas Ann 101, Nr. 8/9, 1977.
- H.-P. BECK: Aktive Schwingungsbedämpfung in Triebachsen mittels selbststellender Zustandsregelung, Elektrische Bahnen 97, 1999.
- JÖCKEL, A. AND R. PFEIFFER: Regelungstechnische Bedämpfung der Reibschwingungen im Antriebsstrang von Drehstromtriebfahrzeugen, Elektrische Bahnen 93, 1995.
- SCHREIBER, R., MUNDY, U., MENSSEN, R. AND R. RUEGG: Der Schritt ins nächste Jahrzehnt: Innovative und kundenorientierte Kraftschlussregelung für Lokomotiven, Schweizer Eisenbahn-Revue, Vol 6, 1998.

Addresses: S. MÜLLER, R. KÖGEL ABB Corporate Research Center, Speyerer Str. 4, 69115 Heidelberg, Germany.



RUDOLPH, M.; POPP, K.

## Modellbildung reibungsselbsterregter Bremsenschwingungen

Es werden Erregungs-Mechanismen (fallende Reibkennlinie, geometrische Instabilität, nichtkonservative Lagekräfte) zur Beschreibung des Energieeintrages von der Bremsscheibe in das mechanische System Bremse dargelegt und diskutiert.

Darauf aufbauend wird ein erweitertes Prinzipmodell vorgestellt, das eine reale Schwimmsattelbremse beschreibt. Dabei handelt es sich um ein Mehrkörpersystem, das elastische Moden der Bremsenbauteile berücksichtigt. Anhand einer Sensitivitäts-Analyse des linearen Modells werden Bauteile mit hoher Relevanz für die Schwingungsanregung identifiziert.

### 1. Einführung

Reibungsselbsterregte Bremsenschwingungen mit Frequenzen oberhalb 1000 Hz werden als Bremsenquietschen bezeichnet. Diese technische Erscheinung gewinnt für den Bremszulieferer in der Automobilindustrie aufgrund des zunehmenden Komfortanspruchs der Automobilkunden immer mehr an Bedeutung.

Da die grundlegenden Wirkmechanismen des Quietschens noch nicht eindeutig identifiziert worden sind, beschränkt sich die Bekämpfung der Geräuschentwicklung auf überwiegend empirische Verfahren. Dabei wird eine bestehende Bremsenkonstruktion, ausgehend von aufwendigen experimentellen Untersuchungen, nachträglich modifiziert. Mit den hier dargelegten Arbeiten wird daher angestrebt, den Wirkmechanismus des Quietschens besser zu verstehen und Parametereinflüsse aufzudecken. Dieses Wissen kann wertvolle Hinweise für den Entwurf und die Konstruktion einer geräuscharmen Bremse liefern.

### 2. Erregungs-Mechanismen

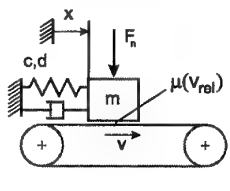
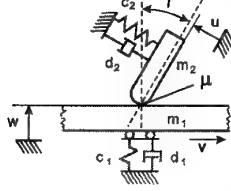
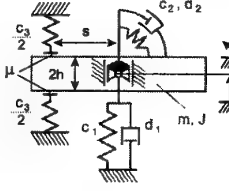
Mechanismus	Reibkennlinie	Geometrische Instabilität	Nichtkons. Lagekräfte
Modell			
Instabilitätsbedingung	$\left. \frac{d\mu}{dv_{rel}} \right _{\dot{x}=0} < -\frac{d}{F_n}$	$\mu > \tan(\gamma) + \frac{2d_2}{\sin(2\gamma)d_1}$	$\mu > \frac{s}{h} + \frac{c_2}{hsc_3} + \frac{c_2}{hsc_1}$

Tabelle 1: Erregungs-Mechanismen des Bremsenquietschens

Im Zustand des Quietschens bildet die Bremse ein selbsterregtes Schwingungssystem, bei dem zwischen Bremsscheibe und Bremsbelag ein Energietransfer stattfindet, der durch die Schwingung gesteuert wird. Die in der einschlägigen Literatur beschriebenen Bremsenmodelle lassen sich auf drei grundlegende Mechanismen dieses Energietransfers zurückführen (Tabelle 1):

**Fallende Reibkennlinie:** Dieser Erklärungsansatz wurde schon 1938 von MILLS in [1] vorgeschlagen. Bedingt durch ein Abfallen des Reibwertes bei zunehmender Relativgeschwindigkeit der Reibpartner kommt es zur Schwingungsanregung bis hin zu Haft-Gleit-Schwingungen (Stick-Slip). Das Modellsystem hierbei ist ein einläufiger Schwinger mit Reibkontakt zu einer bewegten Unterlage. Wird eine Coulombsche Charakteristik angenommen, so verhält sich das System im Großen stabil, Trajektorien ziehen sich auf einen Haft-Gleit-Grenzzykel zusammen. Liegt eine kontinuierlich fallende Kennlinie vor, so bildet die Gleichgewichtslage einen instabilen Fokus, zu erkennen in Abbildung 1a). In Abhängigkeit von den Systemparametern kann das System auch im Großen instabiles Verhalten zeigen.

**Geometrische Instabilität:** Basierend auf experimentellen Erkenntnissen aus der Untersuchung von U-Bahn-Bremsen stellte SPURR 1961 mit [2] eine Erklärungshypothese des Quietschens auf, die von CRISP als mechanisches Modell formuliert wurde. Es handelt sich dabei um einen der Selbsthemmung verwandten Mechanismus, der auf bestimmten geometrischen Verhältnissen und dem Vorhandensein von Reibung im mechanischen System beruht (Sprag-Slip). Die Stabilitätsgrenze wird durch eine Bedingung für den Gleitreibwert bestimmt, die von der Geometrie und der Dämpfung im System abhängt.

**Nichtkonservative Lagekräfte:** Dieses Prinzip findet sich bereits 1972 in den Arbeiten [3] von NORTH und wird hier anhand eines 2-FHG Minimalmodells vorgestellt. Durch reibungsbedingte lageproportionale Kräfte werden Freiheitsgrade schwingungsanregend gekoppelt. Die Anwendung des Hurwitz-Kriteriums liefert mehrere Instabilitätsbedingungen, von denen in Tabelle 1 aus Platzgründen nur eine wiedergegeben ist. Betrachtet man die Systemeigenwerte in Abhängigkeit vom Reibwert, so zeigt sich, dass sich die Eigenwerte der zunächst stabilen Moden mit zunehmenden Reibwert in ihrer Frequenz annähern bevor Destabilisierung des Systems auftritt. Dieser Effekt wird auch als Modenkopplung bezeichnet, vgl. Abbildung 1b).

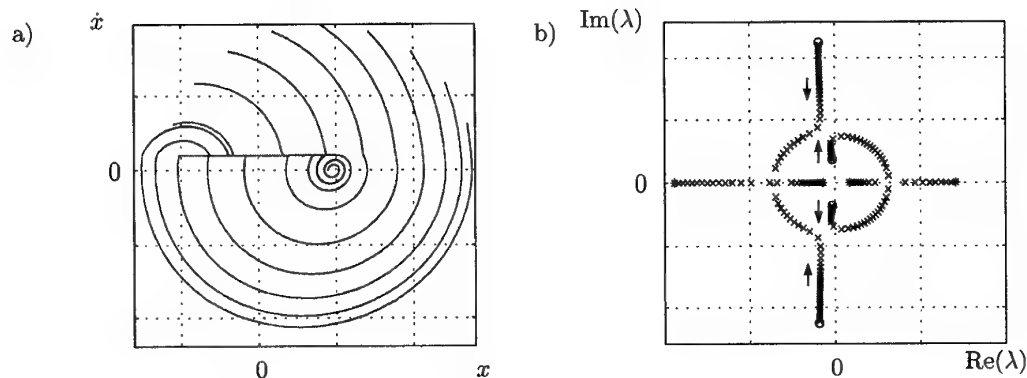


Abbildung 1: Stabilitätsverhalten der Erregungs-Mechanismen

- a) Phasendiagramm zum Reibschwinger mit linear fallender Charakteristik  
 b) Wurzelortskurve des 2FHG-Systems mit nichtkonservativen Lagekräften  
 bei einer Reibwerterhöhung von  $\mu = 0$  (o) auf  $\mu = 1$  (\*).

Um den Mechanismus zu finden, der zum Bremsenquietschen führt, muss die Modelldynamik mit den Eigenschaften des realen Systems verglichen werden.

**Fallende Reibkennlinie:** Nimmt man einen Haft-Gleit-Grenzzykel als Erklärung des Quietschens an, so sollte die Grenzzykel-Periodendauer unabhängig von der Unterlagengeschwindigkeit sein. Diese Forderung ergibt sich aus der Tatsache, dass sich die Quietschfrequenz während der Bremsung nicht kontinuierlich ändert. Der Reibschwinger mit Coulombscher Charakteristik kann dieses Verhalten nicht abbilden; wird aber eine linear fallende Reibkennlinie vorausgesetzt, so lässt sich die Konstanz der Grenzzyklfrequenz bei Veränderung der Unterlagengeschwindigkeit zeigen.

Es konnte jedoch noch nicht experimentell nachgewiesen werden, dass während des Quietschens ein Haften zwischen Belag und Scheibe auftritt.

**Geometrische Instabilität:** Die aus diesem Modell folgende Reibwertbedingung für das Auftreten von selbsterregten Schwingungen, wie sie in Tabelle 1 angegeben ist, erfordert für realistische Stabilitätsgrenzen des Reibwertes, dass die Dämpfung der Scheibe größer als die des Belages ist. Im realen System liegt die Belagdämpfung jedoch gewöhnlich über der Dämpfung der Brems Scheibe.

**Nichtkonservative Lagekräfte:** Dieser Mechanismus kommt ohne die Annahme eines Haftens zwischen Brems Scheibe und Belag aus und erfordert keine Annahmen über die Größe der Dämpfungen im System. Weiterhin lässt sich neben der Wirkung der Steifigkeiten auch der Einfluss von geometrischen Verhältnissen auf die Stabilität erkennen, siehe Tabelle 1. Somit hat der Effekt der nichtkonservativen Lagekräfte einen allgemeinere Bedeutung als die geometrische Instabilität.

Als Konsequenz der hier dargelegten Untersuchung erscheint der Mechanismus der nichtkonservativen Lagekräfte am geeignetsten zur Beschreibung des Energieeintrages beim Bremsenquietschen zu sein.

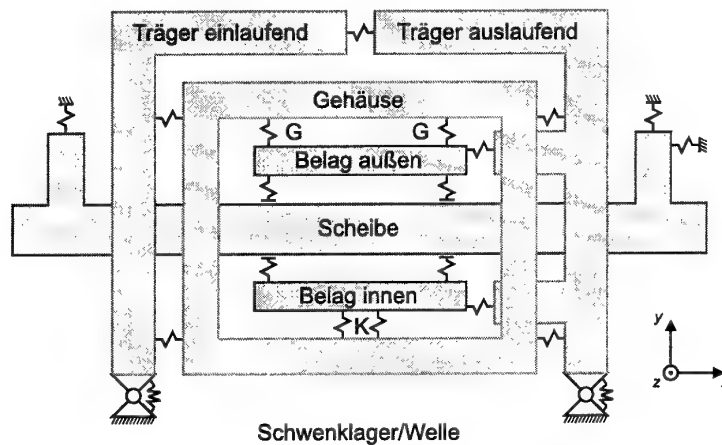


Abbildung 2: Erweitertes Prinzipmodell  
 $G$  Gehäusefinger,  $K$  Kolben

### 3. Erweitertes Prinzipmodell

Um einen Beitrag zur Vermeidung des des Bremsenquietschens leisten zu können, ist es erforderlich, konstruktive Design-Parameter einer Bremse im Modell abzubilden. Umfangreiche Finite-Element-Modelle bieten diese Möglichkeit, besitzen jedoch auch Nachteile. Zum einen ist die Umsetzung von Design-Variationen im Modell aufwendig und zum anderen ergeben sich durch die hohe Anzahl von Freiheitsgraden lange Rechenzeiten. Ausserdem gestaltet sich die Einbindung von harten Nichtlinearitäten schwierig und die gewonnenen Erkenntnisse sind sehr produktspezifisch. Vor diesem Hintergrund wurde quasi als Bindeglied zwischen dem Erregungs-Mechanismus und den Finite-Element-Modellen ein Modell einer Schwimmsattelbremse entwickelt, dass mit wenigen Freiheitsgraden wesentliche Design-Parameter abbildet.

Es handelt sich dabei um ein ebenes Mehrkörpersystem mit 14 Freiheitsgraden, siehe Abbildung 2. Die Parameter dieses Modells wurden zum Teil aus Messungen an den Bremsenbauteilen gewonnen und zum Teil aus einem bestehenden Finite-Element-Modell kondensiert. Elastische Schwingformen sind durch Strukturanpassung beim Bremsenträger und bei der Brems Scheibe durch Parameteranpassung an quietschrelevante Bauteil-Eigenfrequenzen integriert worden. Die Starrkörper des Modells sind verbunden durch lineare Feder-Dämpfer-Koppelemente, die aber durch nichtlineare Elemente ersetzen lassen. In den Koppelementen zwischen Belag und Scheibe ist trockene Reibung integriert, was zu einer Unsymmetrie in der Steifigkeitsmatrix des Systems führt. Das erweiterte Prinzipmodell setzt also auf dem Mechanismus der nichtkonservativen Lagekräfte auf.

Den Kern des Simulationsprogrammes bildet ein allgemeiner Mehrkörper-Algorithmus für Schwingungssysteme mit kleinen Amplituden basierend auf der synthetischen Methode nach Newton und Euler. Die Struktur des zu simulierenden Systems sowie dessen Parameter sind in externen Dateien abgelegt, so dass eine Strukturmodifikation der Bremse einfach in das Modell zu übernehmen ist.

### 4. Technische Sensitivitätsanalyse

Die Stabilität des Systems, und damit die Gräuschneigung der Bremse, wird im linearen Fall durch die Realteile der Eigenwerte beschrieben. Daher erscheint es sinnvoll, die Sensitivität der Realteile bezüglich einer Parametervariation zu untersuchen. FRANKE zeigt in [4], dass sich die Sensitivität eines Systemeigenwertes  $\lambda$  bezüglich eines Parameters  $p$  mit Hilfe der Rechts- und Linkseigenvektoren der Systemmatrix ergibt. Für eine technische Parameter-Variation  $\Delta p$  verändern sich jedoch auch die Systemeigenvektoren um ein signifikantes Maß, so dass die resultierende Veränderung der Eigenwerte nicht mit dem beschriebenen Verfahren zu bestimmen ist. Andererseits stellt es mit der heutigen Rechentechnik kein größeres Problem dar, die Systemeigenwerte vor und nach der Variation  $\Delta p$  zu bestimmen und die Differenz  $\Delta \lambda$  direkt zu berechnen. Für dieses Vorgehen wird der Begriff *technische Sensitivitätsanalyse* eingeführt. Da die Anzahl der Parameter recht groß ist und die Parametervariationen, die untersucht wurden, eine gut zu interpretierende konstruktive Bedeutung haben sollen, wurde die Menge der Parameter in Variationsgruppen eingeteilt, innerhalb derer die Parameter gleichzeitig verändert werden. Die technische Sensitivitätsanalyse wurde zur Untersuchung von zwei Eigenschaften, der Robustheit und der Parametereinflüsse, eingesetzt.

Da der Reibwert zwischen Belag und Scheibe im Bremsbetrieb verhältnismäßig großen Schwankungen unterliegt und ein hoher Reibwert eher zu Geräuschen führt als ein niedriger, wurde die *Robustheit* des Modells gegenüber

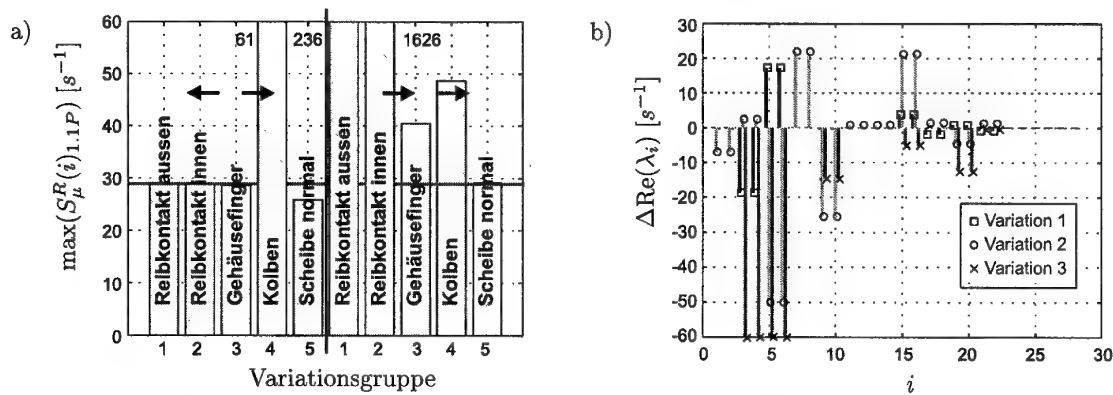


Abbildung 3: Ergebnisse der Sensitivitätsanalyse

a) Robustheitsuntersuchung für gleich- ( $\rightarrow\rightarrow$ ) und gegensinnige ( $\leftarrow\leftarrow$ ) Koppelementverschiebung  
 b) Parametereinfluss

Reibwerterhöhung für etwa einhundert Parametervariationen untersucht. Von den Reibwertsensitivitäten

$$S_{\mu}^R(i) = \frac{\Delta \text{Re}(\lambda_i)}{\Delta \mu} = \frac{\text{Re}(\lambda_i(\mu_2)) - \text{Re}(\lambda_i(\mu_1))}{\mu_2 - \mu_1} \quad (1)$$

aller Eigenwerte wurde jeweils das Maximum ausgewählt und mit dem Maximum des unmodifizierten Zustandes verglichen. Abbildung 3a) zeigt als Ergebnisbeispiel die Wirkung von Koppelementverschiebungen. Es wird deutlich, dass die Steifigkeiten im Reibkontakt und die Verhältnisse zwischen Belag-Rückenplatte und Kolben, bzw. Gehäusefinger, eine grosse Bedeutung haben. In der Praxis werden hier Dämpfungsbleche zur Geräuschminderung eingebaut.

Bei der Untersuchung von *Parametereinflüssen* wurde der Reibwert konstant gehalten und die Parameter entsprechend ihrer Gruppeneinteilung um zehn Prozent variiert. Die resultierende Veränderung der Eigenwert-Realteile ist für alle Eigenwerte einzeln betrachtet worden. Eine Normierung auf den Zahlenwert der Parametervariation wurde zugunsten einer besseren Vergleichbarkeit bei unterschiedlicher Parameter-Dimension nicht vorgenommen. In Abbildung 3b) lässt sich erkennen, dass sich Parametervariationen sehr unterschiedlich im Spektrum der Eigenwerte auswirken können. Die geometrische Variation 3 liefert beispielsweise eine globale Stabilisierung des Systems.

## 5. Zusammenfassung und Ausblick

Die Diskussion der Erregungs-Mechanismen zur Erklärung des Bremsenquietschens zeigt, dass der Mechanismus der nichtkonservativen Lagekräfte einen vielversprechenden Ansatz darstellt. Darauf aufbauend wurde ein Mehrkörper-Modell einer Schwimmsattelbremse entwickelt und einer Sensitivitätsanalyse des Stabilitätsverhaltens bezüglich Parametervariation unterzogen. Die Ergebnisse dieser Sensitivitätsanalyse werden derzeit experimentell überprüft. Weitere Modellrechnungen beziehen nichtlineare Effekte ein, ausserdem wird das Modell auf elastische Körper erweitert.

## Danksagung

Die Arbeiten wurden gefördert durch das Bundesministerium für Bildung und Forschung, FKZ: 13N7202/3, und profitierten von der Zusammenarbeit mit TRW Automotive. Ein besonderer Dank gilt Herrn Dr.-Ing. T. Treyde und Herrn Dipl.-Ing. J. Korte.

## 6. Literatur

- 1 MILLS, H. R.: Brake Squeak. Inst. of Automobile Engineers Report No. 9000 B, 9162 B, (1938,1939).
- 2 SPURR, R. T.: A Theory of Brake Squeal. Proc. Instn. Mech. Engrs. 1 (1961),33-40.
- 3 NORTH, M. R.: Disc Brake Squeal - A Theoretical Model. MIRA Report No. 1972/5, (1972).
- 4 FRANK, P. M.: Introduction to System Sensitivity Theory. Academic Press, New York, San Francisco, London (1978).

Anschrift: M. RUDOLPH, K. POPP, Universität Hannover, Institut für Mechanik, Appelstr. 11, 30167 Hannover, E-Mail: rudolph@ifm.uni-hannover.de

SCHMIEG, H. UND VIELSACK, P.

## Selbsterregte Reibschwingungen bei konstant verzögertem Vortrieb am Beispiel einer Scheibenbremse

*Das Paradigma selbsterregter Reibschwingungen ist ein starrer Körper auf einer rauhen Unterlage, der durch eine Feder mit konstanter Vortriebsgeschwindigkeit geschoben wird. Die Existenz der selbsterregten Schwingungen werden durch die Eigenschaften des Reibgesetzes bestimmt. Im Gegensatz dazu wird für ein erweitertes mechanisches System gezeigt, daß bei konstant verzögertem Vortrieb die Art des Reibgesetzes eine untergeordnete Rolle spielt und die Existenz von instationären selbsterregten Schwingungen im wesentlichen durch die mechanischen Eigenschaften des Systems bestimmt sind.*

### 1. Einleitung

Selbsterregte Reibschwingungen, wie sie bei Scheibenbremsen auftreten, können u.a. durch ihre Frequenzbereiche klassifiziert werden. Höherfrequente Schwingungen ( $2 \text{ kHz} < f < 12 \text{ kHz}$ ) können der Kategorie Bremsenquietschen [1], niederfrequente ( $f < 1 \text{ kHz}$ ) der Kategorie Bremsenrubbeln zugeordnet werden. Ihre Modellierung und der Bereich der Reibkennlinie, welcher entscheidend ist für das Auftreten von selbsterregten Schwingungen, unterscheiden sich wesentlich. In dieser Arbeit wird ausschließlich der Fall niederfrequenter Reibschwingungen betrachtet. Diese sind durch Stick-Slip-Phänomene gekennzeichnet.

### 2. Modellierung und theoretische Untersuchung einer Scheibenbremse

In Bild 1a ist das reale Bremsystem zu sehen, bestehend aus Bremse, Bremsscheibe, Aufhängung und Antrieb. Die Bremse selbst besteht aus dem Bremssattel mit Bremszylinder und dem Träger. Beidseitig der Bremsscheibe liegen im Träger die Bremsbacken (Pads). Beim Bremsen werden diese durch den Bremszylinder gegen die Scheibe gedrückt, von dieser etwas mitgenommen und laufen auf den Träger auf. Die resultierende Reibkraft wird über den an der Aufhängung befestigten Träger in das Fahrzeug eingeleitet.

In dem Versuchsstand Bild 1a wurde die Aufhängung des Fahrzeugs durch einen Biegestab nachgebildet. Der Antrieb erfolgt durch einen drehzahlgeregelten Hydraulikmotor. Im Gegensatz zum realen Fahrzeug ist die Drehzahl der Bremsscheibe unabhängig von der Reibkraft.

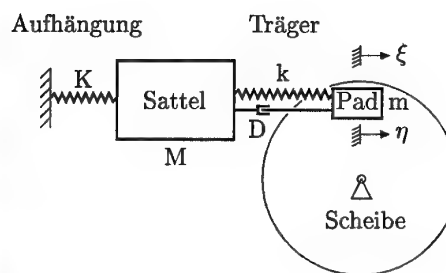
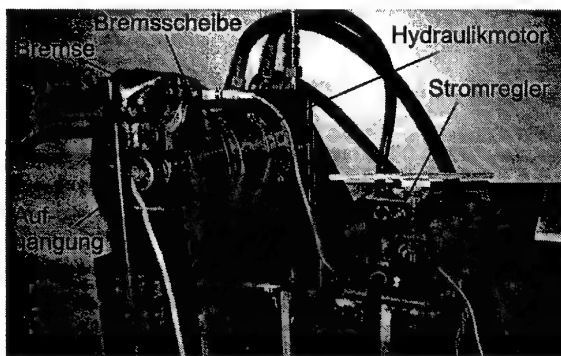


Bild 1

Reales System

Mechanisches Modell

In der Modellierung wurde das Bremsystem durch ein Starrkörpersystem mit diskreten Massen, Federn und Dämpfern, der flächenhafte Reibkontakt Pad/Bremsscheibe durch einen punktförmigen Reibkontakt ersetzt. Allgemein gilt, daß Strukturen größerer Abmessungen beim Schwingen niedrigere, jene mit kleinen Abmessungen höhere Frequenzen abstrahlen. Da hier ausschließlich der niedrige Frequenzbereich betrachtet wird, muß in die Modellierung außer der Bremse auch die umgebende Struktur, also zumindest die Aufhängung, miteinbezogen werden. Bremsenhersteller bestätigen obige Aussage. Basis des mechanischen Modells in Bild 1b ist daher ein 1-Freiheitsgradsystem, bestehend aus der Masse  $M$  der Bremse (im wesentlichen die des Sattels) und der Steifigkeit  $K$  der Aufhängung. Die reibin-

duzierte Anregung erfolgt über ein Pad der Masse  $m$  und die Steifigkeit  $k$  des Trägers. Die eingeführte Dämpfung  $D$  ist im wesentlichen die des Pad. Insgesamt ergibt sich damit als mechanisches Modell ein 2-Freiheitsgradsystem mit Reibung. Dabei wird die Bewegung des Pad durch die Koordinate  $\xi$ , die der Scheibe am Reibkontakt durch die Koordinate  $\eta$  beschrieben.

Das reale Reibgesetz ist durch mechanische Kenngrößen, wie Relativgeschwindigkeit, Anpressdruck, Materialpaarung bestimmt, die wiederum Kurz- und/oder Langzeiteinflüssen, wie Temperatur, Menge und Verteilung des Abriebs, Feuchtigkeit, Belastungsgeschichte, etc., unterliegen. Dies bedeutet, daß das Reibgesetz sich in kaum vorhersagbarer Weise während des Bremsens ändert. Zur theoretischen Berechnung wird das Reibgesetz nach Bild 2 verwendet, welches das reale Gesetz mit all seinen Parametern sicher nur unvollständig erfaßt. Entscheidend für die Untersuchung ist jedoch der Bereich des Reibgesetzes in der Nähe des Ursprungs, da Stick-Slip Phänomene untersucht werden, die wechselnde Haft-Gleitzustände voraussetzen. Experimentell wurden Verläufe dieser Art für die Kombination Padmaterial-Stahl gemessen. Die Reibkraft zeigt eine Haftüberhöhung. Für Gleiten gilt ein lineares Kraftgesetz. Es ist nur von der Relativgeschwindigkeit abhängig; die Normalkraft ist konstant. Für die Berechnung wurden die Werte  $\rho_s = 1$ ,  $\rho_D = 0.75$ ,  $c = -0.02$  angenommen (siehe Bild 2).

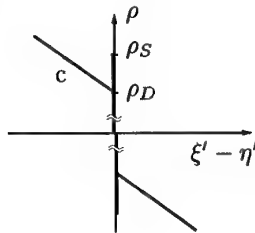


Bild 2: Reibgesetz

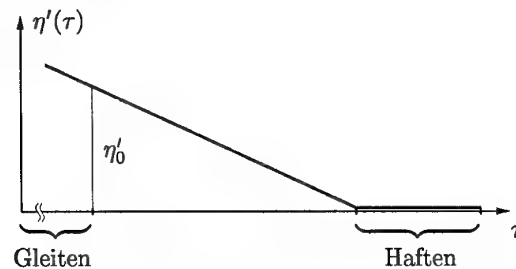


Bild 3: Geschwindigkeitsverlauf  $\eta'(\tau)$  der Scheibe

Eine konstante Verzögerung,  $\eta''$  der Bremsscheibe, bewirkt eine lineare Abnahme der Geschwindigkeit  $\eta'(\tau)$  bis zur Ruhe (Bild 3). Grundzustand ist das permanente Gleiten des Pad auf der Scheibe. Das Pad ist statisch ausgelenkt, seine Geschwindigkeit  $\xi'$  ist zunächst Null. Wird nun bei einer Scheibengeschwindigkeit  $\eta'_0$  eine Störung initiiert, so kann von der gestörten Bewegung des Pads auf die Stabilität dieser Bewegung geschlossen werden. Als Störung wurde ein Haftzustand  $\xi' = \eta'_0$  initiiert. Anhand der Phasenkurve  $\xi'(\xi)$  oder der Kontaktkraft  $\rho(\tau)$  läßt sich leicht entscheiden, ob selbsterregte Reibschwingungen auftreten.

Alle Rechnungen erfolgten in Anlehnung an [2] und, falls nichts anderes angegeben, mit den Werten  $\eta'' = -10^{-3}$ ,  $K/k = 0.15$ ,  $D = 10^{-2}$ ,  $M/m = 10$ .

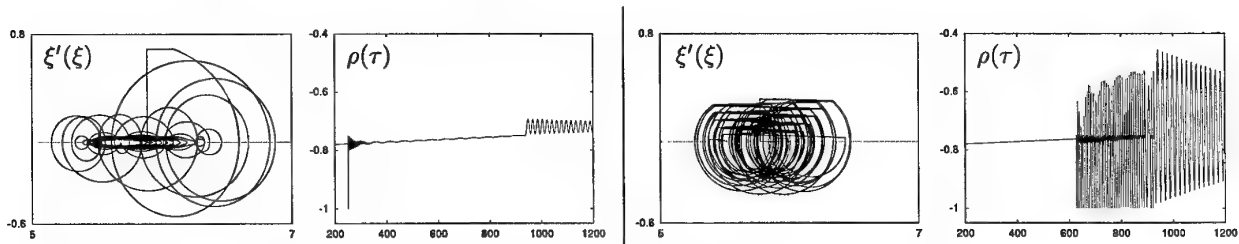


Bild 4: Phasenkurve  $\xi'(\xi)$  und Kontaktkraftverlauf  $\rho(\tau)$  für  $\eta'_0 = 0.7$  und  $\eta'_{0,grenz} = 0.3$

Bilder 4a,b zeigen den Fall  $\eta'_0 = 0.7$ . Ausgehend vom stabilen Zustand  $\xi' = 0$  wird eine Störung  $\xi' = \eta'_0 = 0.7$  initiiert (Vertikale in  $\xi'(\xi)$ ). Diese Haftphase bleibt kurzzeitig erhalten (Horizontale in  $\xi'(\xi)$ ), dann erfolgt der Auslösevorgang. Schon der nächste, kleinere Zyklus zeigt kein Haften mehr. Die Schwingbewegung wird mehr und mehr abgebaut; sie ist stabil. Im Verlauf der Kontaktkraft  $\rho(\tau)$  ist die Störung als abklingende Kurve erkennbar. Danach folgt der Verlauf der Kontaktkraft dem Reibgesetz nach Bild 2 bis zum Stillstand der Scheibe. Der Ausschwingvorgang von  $\rho(\tau)$  resultiert aus der gedämpften Schwingung der Sattelmasse  $M$  bei haftendem Pad auf stillstehender Scheibe. Die Größe dieses Ausschwingvorganges ist von seinen (nichtvorhersagbaren) Anfangsbedingungen abhängig. Stört man das System bei kleineren  $\eta'_0$ -Werten, so erhält man zunächst qualitativ gleiches Verhalten. Erst bei der Grenzgesehwindigkeit  $\eta'_{0,grenz} = 0.3$  ändert sich dieses (Bilder 4c,d). Die Bewegung des Pads wird instabil, d.h. selbsterregte instationäre Reibschwingungen treten bis zum Stillstand der Scheibe auf.

Die Frage ist, welche Größen die Grenzgeschwindigkeit  $\eta'_{0, \text{grenz}}$  beeinflussen. Abgesehen von den Kenngrößen  $M$ ,  $m$ ,  $K$ ,  $k$  und  $D$  des Systems, werden auch die Verzögerung  $\eta''$  der Scheibe und das Kraftgesetz mehr oder weniger dafür entscheidend sein, ob bei höheren oder niedrigen Geschwindigkeiten  $\eta'_0$  Reibschwingungen auftreten. Im folgenden wird die Verzögerung  $\eta''$ , das Steifigkeitsverhältnis  $K/k$  und die Dämpfung variiert. Anhand der Kontaktkraft  $\rho(\tau)$  wird über die Stabilität der Padbewegung befunden (Bild 5).

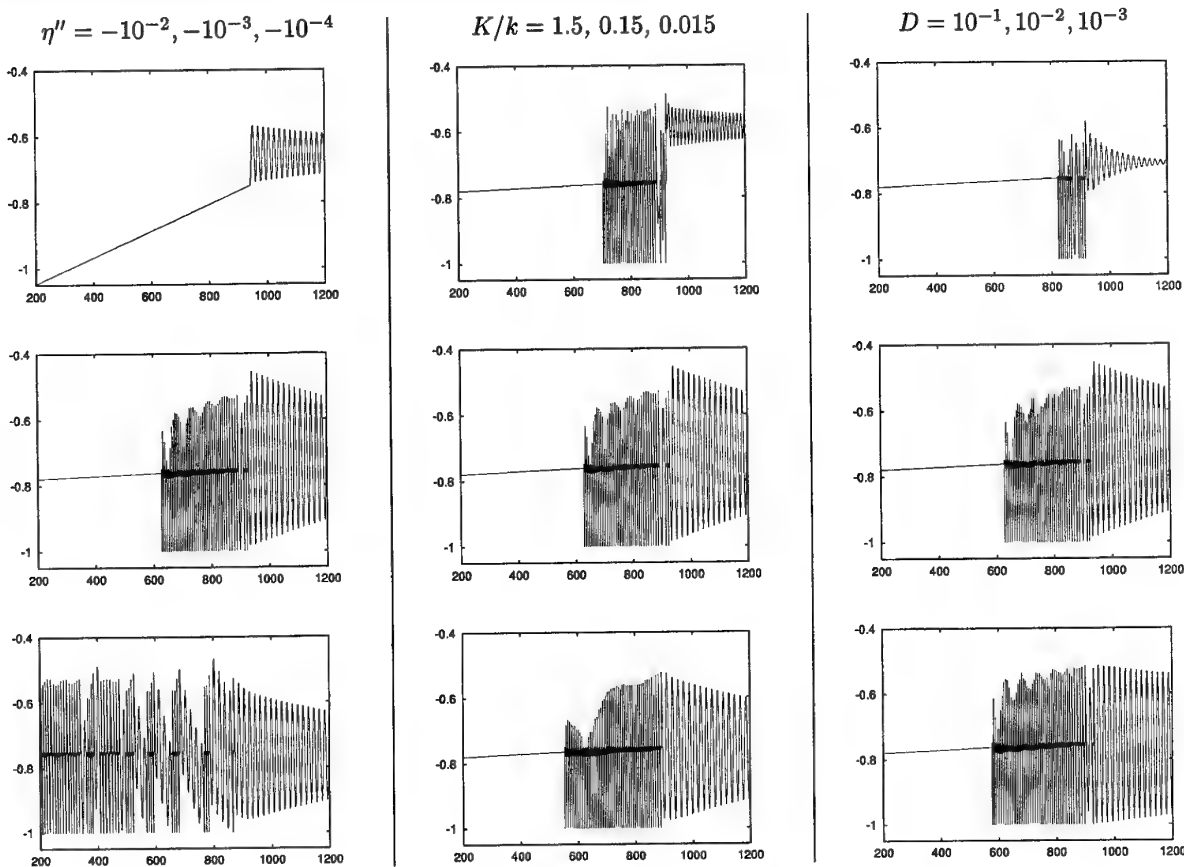


Bild 5: Kontaktkräfte  $\rho(\tau)$  bei Variation der Verzögerung  $\eta''$ , der Steifigkeiten  $K/k$  und der Dämpfung  $D$

Bild 5a ist zu entnehmen, daß mit abnehmender Verzögerung zeitlich früher, d.h. bei höheren Geschwindigkeiten selbsterregte Reibschwingungen auftreten. Verändert man bei gleicher Bremse ( $k = \text{const}$ ) die Steifigkeit  $K$  der Aufhängung, so wird bei weicherer Aufhängung das Pad mehr zum Schwingen neigen (Bild 5b). Wie zu erwarten, ist dies auch der Fall, wenn die Dämpfung  $D$  verkleinert wird (Bild 5c).

Bisher wurde mit dem Reibgesetz nach Bild 2 gerechnet. In [2] wurde gezeigt, daß Systeme mit verzögertem Vortrieb auch für weitere Reibgesetze  $\rho_D \leq \rho_S$ ,  $c \geq 0$  (s. Bild 2) Stick-Slip-Phänomene aufweisen. Dort wurde festgestellt, daß die Existenz von Stick-Slip-Erscheinungen weniger von der Form der verschiedenen Reibcharakteristiken, als von der Eigenschaft des Antriebs abhängt.

### 3. Experimentelle Untersuchungen

Die theoretischen Ergebnisse nach Bild 5 wurden mittels Experimenten verifiziert. Stichprobenartig wurden dazu, bei linear abnehmender Drehzahl  $n$  der Bremsscheibe ( $0 \leq n \leq 1$  U/sec), jeweils drei Beschleunigungs-Zeit-Signale  $a(t)$  des äußeren Pad erfaßt. Variiert wurde die Geschwindigkeitsabnahme, d.h. die Zeit  $T$  bis zum Stillstand, die Steifigkeit  $K$  der Aufhängung und die Dämpfung  $D$ .

In Bild 6 sind die Ergebnisse spaltenweise in 2 Gruppen mit je 3 Bildern dargestellt. Der Gleitzustand (oberstes Bild jeder Gruppe) ist gekennzeichnet durch einen unregelmäßigen  $a(t)$ -Verlauf ohne erkennbares Muster. Dies bedeutet, daß in der Realität permanente Störungen vorhanden sind.



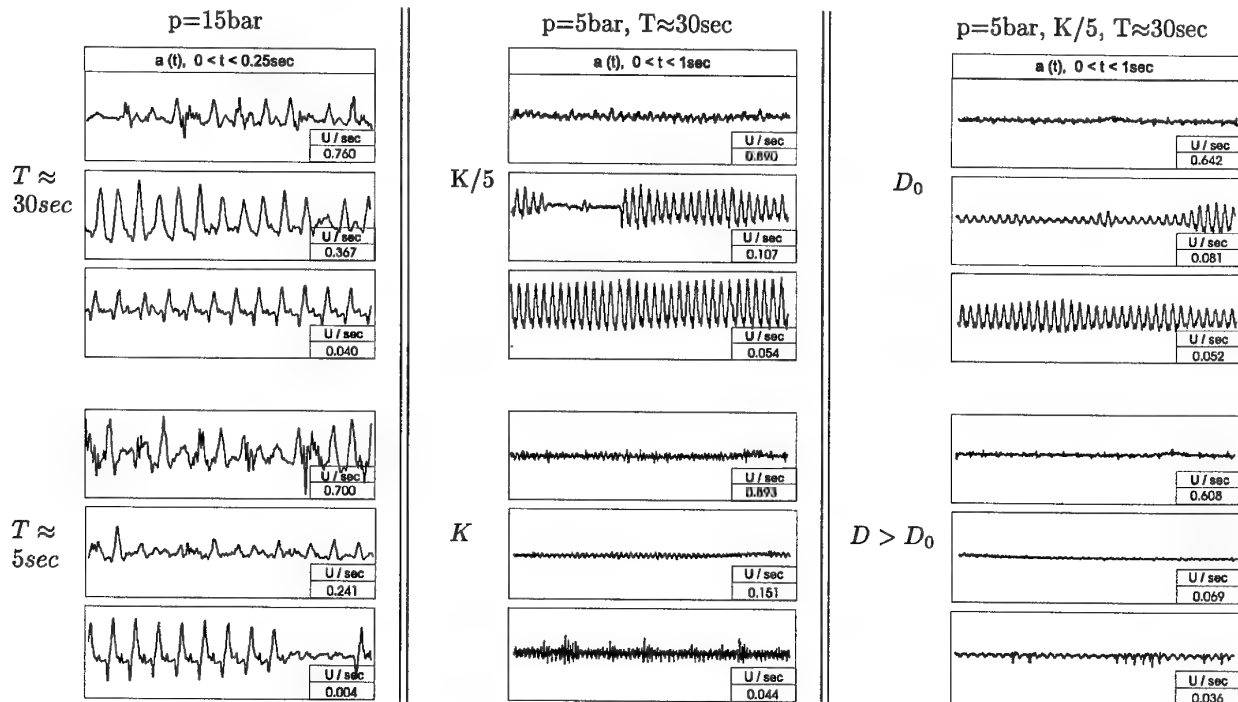


Bild 6: Beschleunigungs-Zeit-Signale  $a(t)$  bei Variation der Bremszeit  $T$ , der Steifigkeit  $K$  und Dämpfung  $D$

In Bild 4a zeigt die obere Gruppe  $a(t)$ -Signale für eine lineare Drehzahlabnahme von 1 U/sec bis zum Stillstand in  $T \approx 30\text{sec}$ , die untere in  $T \approx 5\text{sec}$ . Der oberen Gruppe ist zu entnehmen, daß bei der Drehzahl 0.367 U/sec sich eine regelmäßige Schwingung anfängt auszubilden. Bei niedrigerer Drehzahl erhält man Reibschwingungen mit einer Frequenz von ca. 50Hz, deren Schwingungsamplitude ca. 0.4mm beträgt. Bei schnellem Abfall der Drehzahl bildet sich erst kurz vor dem Stillstand dieses Muster aus. In Übereinstimmung mit Bild 5a ist festzustellen, daß mit abnehmender Verzögerung bei höheren Drehzahlen Reibschwingungen auftreten.

Verringert man den Bremsdruck von  $p=15\text{bar}$  in Bild 6a auf  $p=5\text{bar}$  in Bild 6b, so treten keine Reibschwingungen auf, wie der unteren Gruppe 6b zu entnehmen ist. Wird nun die Steifigkeit  $K$  der Aufhängung auf  $K/5$  verringert (Gruppe 6b, oben), so zeigen sich erneut Reibschwingungen mit, auf Grund der verringerten Steifigkeit, kleinerer Frequenz von ca. 34Hz. Dieses Resultat bestätigt die theoretischen Ergebnisse von Bild 5b.

Die systemimmanente Dämpfung  $D = D_0$  ist im wesentlichen durch das Pad bestimmt. Eine Dämpfungsvergrößerung erfolgte durch Bekleben der Rückseiten der Pads mit Dämpfungsmaterial. Die Meßergebnisse der oberen Gruppe 6c erfolgten mit der Ausgangsdämpfung  $D_0$  und sind eine Wiederholung des Meßvorgangs von Bild 6b, oben. Die untere Bildgruppe 6c zeigt, daß bei größerer Dämpfung Reibschwingungen ganz unterdrückt werden können. Diese Ergebnisse bestätigen die theoretischen Resultate von Bild 5c.

#### 4. Zusammenfassung

Insgesamt ergab sich eine gute Übereinstimmung von Theorie und Experiment in phänomenologischer Hinsicht, was die Hinzunahme der Aufhängung in das Modell bestätigt. Ein verzögerter Vortrieb verstärkt die Tendenz zur Selbsterregung im Vergleich zu jenem mit konstanter Drehzahl. Seine Eigenschaften bestimmen im wesentlichen die Existenz von Stick-Slip-Phänomenen und weniger die Form der Reibungscharakteristik.

#### 5. Literatur

- SCHMIEG, H.; VIELSACK, P.: Modellbildung und experimentelle Untersuchungen zum Bremsenquietschen, ZAMM 78(1998) Suppl.2, 709-710
- VIELSACK, P.: Stick-Slip Instability of Decelerative Sliding, Journal of Non-Linear Mechanics (accepted for publication)

Adresse: DR.-ING. H. SCHMIEG, PROF. DR.-ING. P. VIELSACK, Universität Karlsruhe, Institut für Mechanik, Kaiserstr. 12, 76128 Karlsruhe, Deutschland. E-mail: Mechanik@bau-verm.uni-karlsruhe.de



GERSTEN, K.

## Asymptotic Theory for Turbulent Shear Flows at High Reynolds Numbers

*It is shown that a complete asymptotic theory of turbulent shear flows at high Reynolds numbers near walls exists for the following three standard classes of flows: attached boundary layers, Stratford flows ( $\tau_w = 0$ ) and natural convection flows. These flows are characterized by a finite thickness and a layer structure. The Reynolds-averaged Navier-Stokes equations together with an appropriate turbulence model can be solved by the method of matched asymptotic expansions. Hereby the matching conditions between the different layers yield boundary conditions for the solutions of the equations of motion and furthermore conditions, which asymptotically correct turbulence models have to satisfy. As typical results of the asymptotic theory general explicit formulae for the distributions of the shear stress and the heat flux at the wall exist (usually power laws, except the logarithmic laws for attached boundary layers). For more general classes of flow, e.g. boundary layers with separation, combined natural and forced convections, a complete asymptotic theory is not yet available, because their solutions depend on additional coupling parameters that contain the viscosity.*

### 1. Introduction

There is general agreement that Prandtl's boundary-layer theory for laminar flows is an asymptotic theory for the solutions of the full Navier-Stokes equations at high Reynolds numbers. The question arises whether an asymptotic theory exists also for turbulent flows, that means for the solutions of the Reynolds-averaged Navier-Stokes equation. It should be mentioned that a survey on the asymptotic theory for *free* turbulent shear flows has been given in [1]. Hence, in the following only *wall-bounded* turbulent flows will be considered with the restriction of incompressible two-dimensional flows. It was shown in [2] that L. PRANDTL considered the turbulent boundary layer theory as an asymptotic theory for high Reynolds numbers, although he did not use this terminology. All his contributions to attached turbulent boundary layers show the characteristics of an asymptotic theory. In particular, his new system of formulas for turbulent boundary layers published in 1945 [3] is free from the viscosity. It turns out that attached turbulent boundary layers are not the only class of flows where an asymptotic theory is available today. There are two more classes of flows, as will be shown in the following.

### 2. Layer Structure

In contrast to laminar boundary layers turbulent boundary layers have a finite thickness [4], [5], [6]. They also show a layer structure as sketched in Fig. 1, cf. [7], [8]. There are three layers: the fully turbulent layer (the viscosity can be neglected compared to the eddy viscosity), the viscous sublayer (its thickness is small compared to  $\delta$  and the inertia terms can be neglected) and the viscous superlayer (equilibrium of convection and diffusion). It turns out that two viscous layers have universal local solutions when the local values of the wall shear stress and the wall heat flux for the viscous sublayer and the so-called entrainment velocity for the viscous superlayer are known. Since the equations of motion for the fully turbulent layer is free from the viscosity, only *one* boundary-layer calculation is necessary for *all* Reynolds numbers, which is typical for an asymptotic theory. The boundary conditions that the solutions have to satisfy follow from the matching with the outer flow as well as with the viscous sublayer. The latter matching is of central importance for formulating the boundary conditions for  $\eta = y/\delta \rightarrow 0$ .

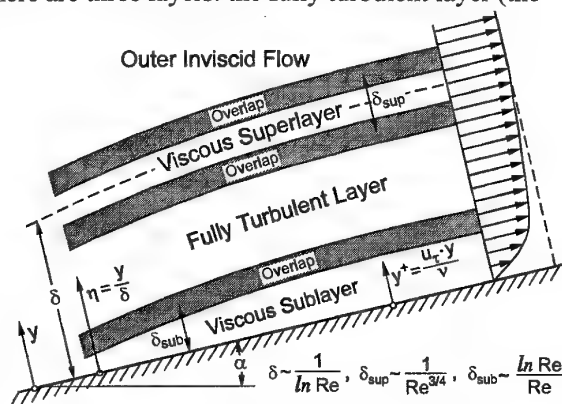


Fig. 1: Attached Turbulent Boundary Layer

### 3. Overlap Layer

An overlap layer is characterized by the fact, that it has the properties of both the neighbouring layers. In the case of the overlap layer between the fully turbulent layer and the viscous sublayer its flow is independent of  $v$  (as part of the fully turbulent layer) and independent of  $\delta$  (as part of the viscous sublayer). These features lead via dimensional analysis to a priori statements about the gradients of velocity and temperature. They are given as follows:

$$\lim_{\eta \rightarrow 0} \frac{\eta}{\sqrt{\tau_t / \rho}} \cdot \frac{\partial u}{\partial \eta} = \frac{1}{\kappa}, \quad \lim_{\eta \rightarrow 0} \frac{\eta \cdot \sqrt{\tau_t / \rho}}{-q_t / (\rho \cdot c_p)} \cdot \frac{\partial T}{\partial \eta} = \frac{1}{\kappa_\theta} \quad (1)$$

Hence, the turbulent Prandtl number is always a constant in the overlap layer between the fully turbulent layer and the viscous sublayer:

$$Pr_t = - \lim_{\eta \rightarrow 0} \frac{\tau_t}{q_t / c_p} \cdot \frac{\partial T / \partial \eta}{\partial u / \partial \eta} = \frac{\kappa}{\kappa_\theta} \quad (2)$$

The balance equations for the flow in this overlap layer reduce to:

$$\text{momentum:} \quad 0 = g \cdot \beta \cdot (T - T_w) \cdot \sin \alpha - \frac{1}{\rho} \cdot \frac{dp}{dx} + \frac{\partial}{\partial y} \left( \frac{\tau_t}{\rho} \right) \quad (3)$$

$$\text{thermal energy (no dissipation):} \quad 0 = \frac{\partial q_t}{\partial y} \quad (4)$$

From this last equation it follows that in the overlap layer the turbulent heat flux is constant and equal to the wall heat flux,  $q_t = q_w$ .

For three classes of flows Eq.(3) reduces to simple relations for the turbulent shear stress as listed in Table 1. The matching conditions (1) lead to the gradients of velocity and temperature in the overlap layer for these three classes of flows listed in Table 1.

These conditions are valid independent of the turbulence model used. On the contrary the constants of the turbulence must be chosen such, that the conditions given in Table 1 are satisfied, cf. [6].

Flow	Definition	$\tau_t$	$\frac{\partial u}{\partial \eta}$ $\sim \frac{\sqrt{\tau_t / \rho}}{\eta}$	$\frac{\partial T}{\partial \eta}$ $\sim \frac{q_t / (\rho \cdot c_p)}{\eta \sqrt{\tau_t / \rho}}$	Constants	$Pr_t$
Attached Boundary Layers	$g = 0$ $\tau_w \neq 0$	$\tau_w$	$\frac{1}{\kappa \cdot \eta}$	$\frac{1}{\kappa_\theta \cdot \eta}$	$\kappa = 0.41$ $\kappa_\theta = 0.46$	$\frac{\kappa}{\kappa_\theta} = 0.89$
Stratford Flows	$g = 0$ $\tau_w = 0$	$\sim \frac{dp}{dx} \cdot \eta$	$\frac{1}{\kappa_\infty \cdot \eta^{1/2}}$	$\frac{1}{\kappa_{\infty\theta} \cdot \eta^{3/2}}$	$\kappa_\infty = 0.59$ $\kappa_{\infty\theta} = 0.72$	$\frac{\kappa_\infty}{\kappa_{\infty\theta}} = 0.82$
Natural Convection	$g \neq 0$ $\frac{dp}{dx} = 0$	$\sim \eta^{2/3}$	$\frac{1}{\kappa_N \cdot \eta^{2/3}}$	$\frac{1}{\kappa_{N\theta} \cdot \eta^{4/3}}$	$\kappa_N = 0.32$ $\kappa_{N\theta} = 0.18$	$\frac{\kappa_N}{\kappa_{N\theta}} = 1.8$

Table 1: Gradients of Velocity and Temperature in Overlap Layer

#### 4. Attached Boundary Layers

The equations of motions combined with appropriate turbulence-model equations are free from the viscosity for the fully turbulent layer. The boundary conditions for  $\eta = 1$  are given by matching with the outer flow and they follow for  $\eta \rightarrow 0$  from the conditions in Table 1. Only *one* calculation is necessary for all Reynolds numbers, cf. [4]. The main result of this calculation is given by the functions  $\tilde{C}(x)$  and  $\tilde{C}_\theta(x)$ . Matching the velocities of the fully turbulent layer and the viscous sublayer leads to the analytical formula for the local skin-friction coefficient  $c_f = 2 \tau_w / (\rho U^2)$ :

$$\sqrt{\frac{2}{c_f}} = \frac{1}{\kappa} \cdot \ln \left( \text{Re}_x \cdot \frac{c_f}{2} \right) + C^+ + \tilde{C}(x) \quad (5)$$

Where  $\text{Re}_x = U(x) \cdot [x - x_0] / \nu$  is the local Reynolds number and  $C^+$  a universal constant ( $C^+ = 5.0$  for smooth surfaces). The main contribution of the right hand side of Eq. (5) comes from the first two terms. They are results of the universal solution for the viscous sublayer. The function  $\tilde{C}(x)$  is the contribution of the fully turbulent layer and hence dependent on the turbulence model. Its effect on the skin friction, however, is only a few percent and decreases with growing Reynolds number.

It is worth mentioning that the so-called *equilibrium boundary-layers* are characterized by a constant  $\tilde{C}$ , in which case the profiles of the velocity defect in the fully turbulent layer are self-similar and the partial differential equation reduces to an ordinary differential equation.

A formula similar to Eq. (5) is valid for the local wall heat flux or Nusselt number, cf. [8].

#### 5. Stratford Flows

In Stratford flows the wall shear stress is everywhere equal to zero ( $\tau_w = 0$ ). STRATFORD [9] has investigated such a flow experimentally. Theoretical investigations of this particular equilibrium boundary layer can be found in [8], [10], [11]. As a result the following formula for the Stanton number, i.e. the dimensionless local wall heat flux  $q_w(x)$ , can be derived:

$$St = \frac{q_w}{\rho \cdot c_p \cdot (T_w - T_\infty) \cdot U(x)} = \frac{0.6}{C_{\theta\infty}^x(\text{Pr}) \cdot \text{Re}_x^{1/3}} \quad (6)$$

Here  $\text{Re}_x$  is again the local Reynolds number based on the outer flow velocity  $U(x) \sim (x-x_0)^{-0.22}$  and the length  $x - x_0$ . The value  $C_{\theta\infty}^x(\text{Pr})$  results from the universal viscous sublayer solution. It depends on the Prandtl number  $\text{Pr}$ , it is  $C_{\theta\infty}^x(0.72) = 1.8$ , cf. [12], [13].

#### 6. Natural Convection Flows

The equations of motion for the fully turbulent layer can be found in [4], [8]. In these particular cases it is not necessary to solve these equations to get a formula for the wall heat flux. The matching between the fully turbulent layer and the viscous sublayer yields directly the following formula to the leading order:

$$\frac{T_w - T_\infty}{T_q} = C_{N\theta}^x(\text{Pr}) \quad (7)$$

where

$$T_q = \frac{q_w}{\rho \cdot c_p \cdot u_q} \quad u_q = \left( \frac{\nu \cdot q_w \cdot \beta \cdot g \cdot \sin \alpha}{\rho \cdot c_p} \right)^{1/4} \quad (8)$$

The function  $C_{N\theta}^x(\text{Pr})$  is given in [4]. It is  $C_{N\theta}^x(0.72) = 4.56$ . Equation (7) is equivalent to a power law of the Nusselt number as function of the Rayleigh number. The exponent of the power law depends on the input (given  $T_w(x)$  or  $q_w(x)$ ). The agreement of Eq (7) with experiments is excellent, cf. [4].

## 7. Incomplete Asymptotic Theory

The momentum equation, Eq. (3), for the overlap layer in general can be written in dimensionless form as follows

$$\tau_t^+ = \frac{\tau_t}{\rho \cdot u_\tau^2} = 1 + K \cdot y^+ + K_N \cdot \int_0^{y^+} \Theta^+(y^+) \cdot dy^+ \quad (9)$$

where

$$K = \frac{\nu}{u_\tau \cdot \tau_w} \cdot \frac{dp}{dx} \quad K_N = \frac{\nu \cdot q_w \cdot \rho \cdot \beta \cdot g}{c_p \cdot \tau_w^2} \quad (10)$$

are so-called coupling parameters. They represent a distinguished limit for the double limiting process  $\nu \rightarrow 0$ ,  $\tau_w \rightarrow 0$ .

$K$  is a parameter that changes along the wall for turbulent boundary layers with separation, cf. [4], [12]. Their fully turbulent layers depend also on  $K$  and hence on the viscosity. Therefore, a complete asymptotic theory with a solution that is independent of the Reynolds number is not available yet. Consequently, turbulence models that are used to predict turbulent boundary layers with separation must have model constants which are functions of the coupling parameter  $K$ . For  $|K| > 0.01$  the effect of  $K$  on the solution has to be taken into account, cf. [4]. The coupling parameter  $K_N$  serves in analogous form for the combined forced and natural convection flows along a flat plate. In the most general case (pressure gradient as well as buoyancy effects) both coupling parameters may become important.

## 8. References

- 1 SCHNEIDER, W.: Boundary-layer theory of free turbulent shear flows. Z. Flugwiss. Weltraumforsch. Bd. 15 (1991), 143 – 158.
- 2 GERSTEN, K.: Ludwig Prandtl und die asymptotische Theorie für Strömungen bei hohen Reynolds-Zahlen. In: G.E.A. MEIER (Hrsg.): Ludwig Prandtl, ein Führer in der Strömungslehre. Vieweg - Verlag, Braunschweig, Wiesbaden, 2000, 125 – 138.
- 3 PRANDTL, L.: Über ein neues Formelsystem für die ausgebildete Turbulenz. Nachr. Akad. Wiss. Göttingen, Math.-phys. Klasse, 1945, 6 - 19.
- 4 SCHLICHTING, H.; GERSTEN, K.: Boundary Layer Theory, Springer-Verlag, Berlin / Heidelberg, 8. Edition, 2000.
- 5 WILCOX, D.C.: Turbulence Modeling for CFD. DCW Industries Inc., La Canada, California, Second Edition, 1998.
- 6 JEKEN, B.: Asymptotische Analyse ebener turbulenter Strömungen an gekrümmten Wänden bei hohen Reynolds-Zahlen mit einem Reynolds-Spannungs-Modell. VDI-Berichte, Reihe 7, No. 215, VDI-Verlag, Düsseldorf, 1992.
- 7 GERSTEN, K.: Turbulent boundary Layers I: Fundamentals. In: A. KLUWICK (Ed.): Recent Advances in Boundary Layer Theory. CISM Courses and lectures No. 390. Springer Wien, New York, 1998, 107 – 144.
- 8 GERSTEN, K.; HERWIG, H.: Strömungsmechanik. Grundlagen der Impuls-, Wärme- und Stoffübertragung aus asymptotischer Sicht. Vieweg-Verlag, Braunschweig / Wiesbaden, 1992.
- 9 STRATFORD, B.S.: An experimental flow with zero skin friction throughout its region of pressure rise. I. Fluid Mech. Vol. 5, 1959, 17 - 35.
- 10 SZABLEWSKI, W.: Inkompressible turbulente Temperaturgrenzschichten mit konstanter Wandtemperatur. Int. J. Heat and Mass Transfer, Vol. 15, 1972, 673 – 706.
- 11 VIETH, D.: Berechnung der Impuls- und Wärmeübertragung in ebenen turbulenten Strömungen mit Ablösung bei hohen Reynolds-Zahlen. VDI-Fortschritt-Berichte, Reihe 7, Nr. 311, 1997.
- 12 VIETH, D.; KIEL, R.; GERSTEN, K.: Two-dimensional turbulent boundary layers with separation and reattachment including heat transfer. In: M. FIEBIG, N.K. MITRA (Eds.): Vortices and Heat Transfer. Notes on Numerical Fluid Mechanics, Vol. 63, Vieweg-Verlag, Braunschweig, Wiesbaden, 1998, 63 – 103.
- 13 VIETH, D.: Heat and momentum transfer in turbulent boundary layers in the presence of strong adverse pressure gradients. In: GERSTEN, K. (Ed.): Asymptotic Methods for Turbulent Shear Flows at High Reynolds Numbers. Kluwer Academic Publishers, Dordrecht, Boston, London, 1996, 155 – 168.

Address: PROF. DR.-ING. KLAUS GERSTEN, Ruhr-Universität Bochum, Institut für Thermo- und Fluidodynamik, D-44780 Bochum, Germany.

A. KLUWICK

## Laminar boundary layer separation, fast and slow

*It is well known that classical boundary layer theory fails if flow separation occurs. In the limit of large Reynolds number this failure can be avoided if the interaction between the viscous wall layer and the external inviscid region is accounted for. The form of the resulting interaction equations crucially depends on the route towards separation. If a firmly attached boundary layer is separated by the action of a rapid pressure increase separation is governed by the triple deck equations e.g. the classical boundary layer equations supplemented with appropriate matching and interaction conditions. However, if the approach towards separation is much slower the formation of a marginally separated flow region is described by a nonlinear integrodifferential equation. Representative solutions of both sets of interaction equations indicate that boundary layer separation often is accompanied with a loss of uniqueness.*

Without doubt boundary layer theory published 1904 in the seminal paper "Über die Flüssigkeitsbewegung bei sehr kleiner Reibung" by L. Prandtl represents one of the cornerstones of modern fluid mechanics. Nevertheless, one must concede that the classical hierarchical concept, in which the pressure distribution inside the boundary layer is imposed by the external inviscid flow and thus known in advance at each level of approximation, is able to yield complete solutions to given problems in rare cases only. The most well known example is provided by a semi infinite aligned flat plate in a uniform stream of an incompressible fluid. However, difficulties arise if one considers a plate of finite length  $\bar{L}$ . This has been pointed out first by GOLDSTEIN 1930 who showed that both the transverse velocity component and the first order pressure disturbances resulting from the boundary layer displacement exhibit singularities at the trailing edge. Despite the occurrence of these singularities, however, the solution to the boundary layer equations can be extended into the wake region in a sensible manner.

A more severe breakdown of classical boundary layer theory is usually encountered if flow separation occurs. In this connection it is found that two routes towards separation have to be distinguished. This is seen most easily if one considers the flow past a slender airfoil at a small angle of attack. First let us concentrate on the case of trailing edge stall. To bring out the essential features unencumbered by complicated geometry the airfoil - following BROWN AND STEWARTSON 1970 - is replaced by a flat plate. If the Kutta condition is enforced the pressure on the suction side rises to its ambient value as the trailing edge is approached where the adverse pressure gradient is infinite. For the case of laminar flow considered here this means that the boundary separates before the trailing edge is reached even if the angle of attack  $k$  is arbitrarily small. For  $k \ll 1$  the transition from a fully attached to a separated boundary layer occurs very fast, e.g. in the immediate neighbourhood of the trailing edge. It is characterised by the formation of a Goldstein singularity in the wall shear stress-distribution and, most important, it is found that solution of the boundary layer equations cannot be extended beyond the point of vanishing wall shear. A different flow behaviour is observed if one considers the case of leading edge stall, RUBAN 1982, RUBAN 1981, STEWARTSON 1982. As the fluid passing over the suction side of the airfoil accelerates starting at the stagnation point the pressure drops rapidly. This initial phase of flow development is followed by a sharp pressure rise which may cause the wall shear in the nose region to decrease significantly. However, if the angle of attack  $k$  is sufficiently small the wall shear remains positive there and rises again if the distance  $\tilde{x}$  from the leading edge is larger than  $\tilde{x}_s$  say. Further increase of  $k$  eventually leads to the wall shear distribution also shown in Fig. 1: the wall shear  $\tilde{\tau}_w$  vanishes in a single point but immediately recovers. In contrast to the Goldstein singularity  $d\tilde{\tau}_w/d\tilde{x}$  remains finite at  $\tilde{\tau}_w = 0$ . It changes discontinuously and the solution - in which the point of zero wall shear is approached much slower than before - can be continued further downstream. If the angle of attack is increased beyond the critical value  $k_c$  leading to this so-called marginal separation singularity a Goldstein singularity forms upstream of  $\tilde{x}_s$  and, as before, represents an impassé for the boundary layer calculations. This fact should, however, not be misinterpreted as a signal that the boundary layer equations cease to be valid at separation. In fact, they remain valid in both cases of *fast* trailing edge and *slow* leading edge separation. Rather it is the hierarchical structure of classical boundary layer theory which breaks down, namely the assumption that the pressure distribution inside the boundary layer and the solution to the boundary layer equations can be obtained in successive steps. This conclusion was drawn independently by a number of authors including NEILAND 1969, STEWARTSON 1969 and MESSITER 1970. It clearly indicates how boundary layer theory can be modified to allow for the description of separation processes both fast and slow: by accounting for the displacement of the boundary layer on the external inviscid flow in leading rather than higher order or, in other words, by letting the boundary layer and the outer inviscid region to interact.

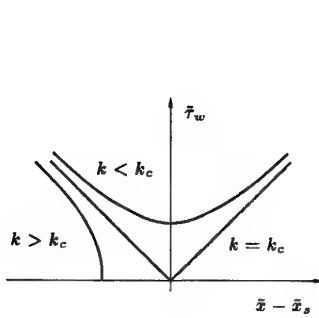


Fig. 1

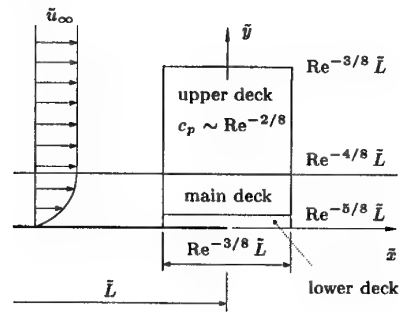


Fig. 2

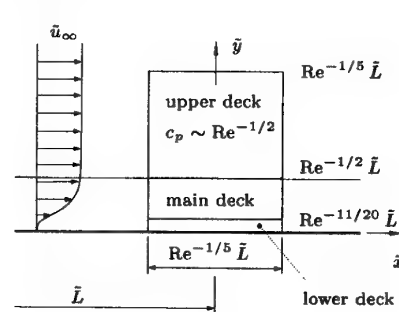


Fig. 3

In the first case one is led to the so-called triple deck theory. Its name derives from the finding that the local interaction region exhibits a three layer structure, Fig. 2. Asymptotic analysis for large Reynolds number  $Re = \bar{u}_\infty \bar{L} / \bar{\nu} \gg 1$  reveals that the length and width of the interaction zone are of the order  $Re^{-3/8} \bar{L}$  where  $\bar{u}_\infty$  and  $\bar{\nu}$  represent reference values of the external flow velocity and the kinematic viscosity while  $\bar{L}$  is a characteristic length associated with the unperturbed boundary layer. Since the interaction length is so short the disturbances generated, for example, by a trailing edge are essentially inviscid not only in the region outside the boundary layer (upper deck) but also over most of the boundary layer (main deck). Viscous effects play a major role in a thin layer (lower deck) adjacent to the wall and the wake centerline. Here the flow is governed by the boundary layer equations. Furthermore, even if the flow under consideration is compressible, the velocities close to the wall are small and the field quantities in the lower deck region are thus governed by the incompressible version of the boundary layer equations. Using suitably scaled coordinates  $(X, Y)$  parallel and normal to the wall, corresponding velocity components  $(U, V)$  and pressure disturbances  $P$  they assume the form

$$\frac{\partial U}{\partial X} + \frac{\partial V}{\partial Y} = 0, \quad U \frac{\partial U}{\partial X} + V \frac{\partial U}{\partial Y} = -\frac{dP}{dX} + \frac{\partial^2 U}{\partial Y^2} \quad (1)$$

The boundary conditions

$$X < 0, Y = 0: \quad U = V = 0, \quad X > 0, Y = 0: \quad \text{continuity of pressure,} \quad (2)$$

$$X \rightarrow -\infty: \quad U = Y, \quad P = -\alpha(-X)^{1/2}$$

include the no-slip condition at the plate, the requirement that the pressure is continuous across the wake centerline and exhibits the upstream behaviour predicted by the theory of inviscid flows. Here  $\alpha = k Re^{1/6}$  denotes the scaled angle of attack. Additional conditions

$$Y \rightarrow \infty: \quad U = Y + A(X), \quad P(X) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{A'(\bar{X})}{X - \bar{X}} d\bar{X} \quad \text{subsonic,} \quad -A'(X) \quad \text{supersonic flow} \quad (3)$$

follow from the requirement that the flow properties in the lower, main and upper deck blend smoothly, e.g. can be matched. Here  $-A(X)$  characterises the displacement effect exerted by the lower deck which is felt (through the passive main deck) by the upper deck where it causes an inviscid pressure response.

The interaction equations 1 - 3 for the suction side of the plate have to be supplemented with similar equations for the pressure side and then have to be solved numerically. Before turning to a discussion of numerical results it is useful to briefly describe the nature of the interaction process if the approach to separation is slow. As before the local interaction region exhibits a three layer structure, Fig. 3. Again, the flow in the main deck region which comprises most of the boundary layer is essentially passive and outside the boundary layer we have a weakly perturbed parallel flow. Viscous effects associated with the interaction process are confined to a thin lower deck region where the boundary layer equations hold. In contrast to the triple problem outlined earlier, however, the incoming flow is no longer fully attached but on the verge of separation. This in turn means that very small disturbances are sufficient to separate the boundary layer and asymptotic analysis indicates that these disturbances satisfy the boundary layer equations which are linearized with respect to the separation profile. The first order solution is seen to include a function  $-A(X)$  which remains undetermined at this level of approximation and, as earlier, can be interpreted as a perturbation displacement thickness or equivalently as a scaled wall shear. In order to determine  $A(X)$  one has to investigate the second order lower deck problem which is found to have an acceptable solution only if  $A$  and the

scaled pressure disturbances  $P$  satisfy the relationship

$$A^2(X) - X^2 + \Gamma = - \int_{-\infty}^X \frac{P'(\bar{X})}{\sqrt{X - \bar{X}}} d\bar{X} \quad (4)$$

where the parameter  $\Gamma$  measures the difference between  $k$  and  $k_c$ :  $\Gamma \propto (k - k_c)Re^{2/5}$ . A second relationship between  $A$  and  $P$  is given by the Hilbert integral or the Ackeret formula in (3) if the external flow is subsonic or supersonic.

Now let us return to the triple deck problem for the flat plate at incidence. Its solutions determine, among others, the dependence of the lift coefficient on the angle of incidence and the Reynolds number:  $c_L/2\pi = 1 - a_1(\alpha)Re^{-3/8}$ . On physical grounds one expects that viscous effects will reduce the lift which requires  $a_1$  to be positive and this is confirmed by the numerical results of CHOW AND MELNIK 1976, KOROLEV 1989, Fig. 4. As one expects the flow remains attached if the angle of incidence is sufficiently small. Most interesting, however, it is found that solutions of the interaction problem do not exist if  $\alpha$  exceeds the critical value  $\alpha_c \approx 0.497$  (indicating that larger values of  $\alpha$  will lead to a substantial change of the flow structure) and that the relationship between  $a_1$  and  $\alpha$  is non unique in the neighbourhood of  $\alpha_c$ . Upper branch solutions exhibit much longer separated flow regions than lower branch solutions which results in a substantial reduction of lift.

The phenomenon of non-uniqueness observed in Fig. 4 appears to occur quite frequently in connection with flow separation. In fact it has been detected first in the different context of supersonic flows past flared cylinders GITTNER AND KLUWICK 1987. The flow field in the neighbourhood of the corner develops the triple-deck structure discussed before with the difference that the interaction law is more complicated than (3). Owing to the axisymmetric geometry the streamtube area decreases downstream of the corner if the flare angle is negative and the initial pressure drop, therefore, is followed by a pressure rise which may separate the boundary layer. The positions  $X_S$  and  $X_R$  of the separation and reattachment points are plotted in Fig. 5 which clearly displays ranges of the (scaled) flare angle  $\alpha$  where the solutions of the interaction problem are non unique.

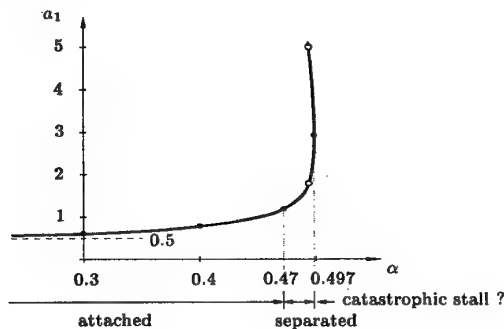


Fig. 4

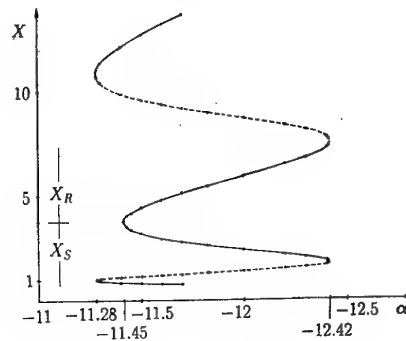


Fig. 5

In the examples discussed so far the external flow was either purely subsonic or supersonic. The treatment of transonic interactions is severely hampered by the nonlinearity of the transonic small perturbation equation which has to be solved simultaneously with the nonlinear boundary layer equations and very limited progress has been achieved so far. Some insight into the properties of mixed subsonic supersonic flow can, however, be gained if one considers transonic effects in narrow channels, KLUWICK AND GITTNER 2001. Owing to the slenderness of the channel the flow in the inviscid core region is essentially one-dimensional and can be investigated analytically. This leads to a nonlinear interaction relationship between  $P$  and  $A$  which allows for the passage through the critical state.

At this point let us stop this brief discussion of fast separation and to return to the case of slow separation. Again, non-uniqueness of flow patterns is found to represent a very common phenomenon. In fact, it occurs in all known examples of marginally separated flows. Results for the problem of leading edge stall, which was used to introduce the notion of marginal separation, are shown in Fig. 6 which displays the wall shear at  $X = 0$  as a function of the parameter  $\Gamma$ . If  $\Gamma$  is negative, i.e. if  $k < k_c$  the flow remains attached and the solutions of the interaction equations 3, 4 are unique. However, for all values  $\Gamma > 0$  for which solutions can be found there exist at least two different ones. Furthermore, it is interesting to note that the interaction equations do not admit solutions if  $\Gamma \gtrsim 2.75$  indicating that a substantial change of the flow behaviour - not fully understood at present - must take place as  $\Gamma$  exceeds this value. Also note, that there exists a  $\Gamma$ -range where one has four rather than two different solutions. There exist cases of marginally separated flows, however, where non-uniqueness of solutions is even more pronounced. An example is



provided by a viscous jet flowing past a curved wall which separates under the action of an adverse pressure gradient generated by centrifugal forces, ZAMETAEV 1986. If the wall shear at  $X = 0$  is plotted as a function of the parameter  $\Gamma$  which characterizes the flow angle downstream of the interaction region (measured from the value where classical boundary layer theory predicts the onset of separation) one obtains the plot shown in Fig. 7. As before there exists an upper limit  $\Gamma_c$  of  $\Gamma$  for which solutions of the interaction problem can be found. For  $\Gamma < \Gamma_c$  there exist ranges of  $\Gamma$  where we have two, four, six or even more different solutions. In this connection the question arises how these solutions react to small disturbances. Disturbances caused by localized surface mounted obstacles  $Y = H(X, Z)$  have recently investigated by BRAUN AND KLUWICK 2000, BRAUN, KLUWICK AND STEINRÜCK 2001 for  $\Gamma_c - \Gamma \ll 1$ . Representative results are depicted in Fig. 8 where  $A_{3D}$  and  $Z$  denote the wall shear distribution in the presence of the 3D obstacle and the (scaled) coordinate in the lateral direction. In the case of the upper branch solutions (u) the disturbances are seen to decrease with increasing distance from the obstacle. Disturbances of the lower branch solutions (l), however, generate a periodic flow pattern which extends up to  $Z = \infty$ . Asymptotic analysis indicates that the flow behaviour for large  $Z$  is almost independent of the specific shape of the obstacle for both the upper and lower branch solutions where they approach the properties of weakly nonlinear eigensolutions. These closely resemble soliton and cnoidal wave solutions known from the Korteweg de Vries equation. The implications of these surprising results are currently under investigation.

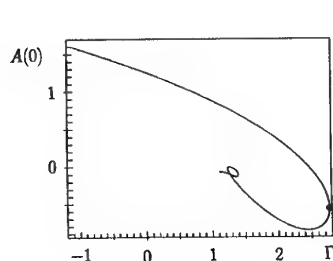


Fig. 6

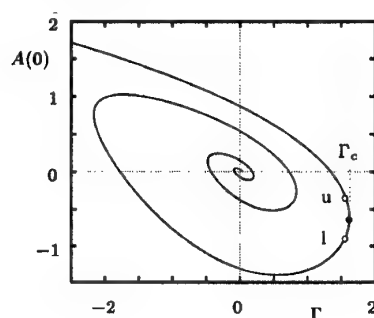


Fig. 7

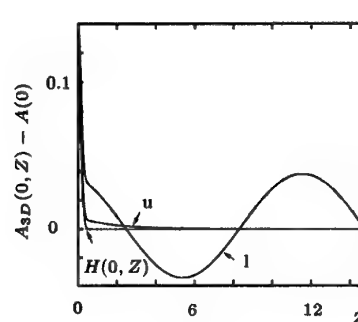


Fig. 8

## References

- 1 BRAUN, S. AND KLUWICK, A.: Two- and three-dimensional marginal separation of laminar, incompressible viscous jets, ZAMM 80, S3 (2000) 623-624.
- 2 BRAUN, S., KLUWICK, A. AND STEINRÜCK, H.: The effect of 3D surface mounted obstacles on laminar, incompressible marginally separated boundary layer flows, ZAMM (2001), submitted.
- 3 CHOW, R. AND MELNIK, R.E.: Numerical solutions of the triple-deck equations for laminar trailing-edge stall, in *Proceedings of the 5th International Conference on Numerical Methods in Fluid Dynamics*, 135-144, Springer, (1976).
- 4 GITTler, P. AND KLUWICK, A.: Triple-deck solutions for supersonic flows past flared cylinders, J. Fluid Mech. 179 (1987) 469-487.
- 5 GITTler, P. AND KLUWICK, A.: Transonic laminar interacting boundary layers in narrow channels, ZAMM (2001), submitted.
- 6 GOLDSTEIN, S.: Concerning some solutions of the boundary layer equations in hydrodynamics, Proc. Camb. Phil. Soc. 26 (1930) 1-30.
- 7 KOROLEV, G.L.: Contribution to the theory of thin profile trailing edge separation, Izv. Akad. Nauk SSSR: Mekh. Zhidk. Gaza 4, 3 (1989) 55-59.
- 8 MESSITER, A.F.: Boundary layer flow near the trailing edge of a flat plate, SIAM J. Appl. Math. 18 (1970) 241-257.
- 9 NEILAND, V.Y.: Towards a theory of separation of the laminar boundary layer in a supersonic stream, Izv. Akad. Nauk SSSR: Mekh. Zhidk. Gaza 4, 4 (1969) 53-57.
- 10 RUBAN, A.I.: Singular solution of boundary layer equations which can be extended continuously through the point of zero surface friction, Izv. Akad. Nauk SSSR: Mekh. Zhidk. Gaza 16, 6 (1981) 835-843.
- 11 RUBAN, A.I.: Asymptotic theory of short separation regions on the leading edge of a slender airfoil, Izv. Akad. Nauk SSSR: Mekh. Zhidk. Gaza 17, 1 (1982) 33-41.
- 12 STEWARTSON, K.: On the flow near the trailing edge of a flat plate II, Mathematika 16 (1969) 106-121.
- 13 BROWN, S.N. AND STEWARTSON, K.: Trailing edge stall, J. Fluid Mech. 42, 2 (1970) 561-584.
- 14 STEWARTSON, K., SMITH, F.T. AND KAUPS, K.: Marginal separation, Stud. Appl. Math. 67 (1982) 45-61.
- 15 ZAMETAEV, V.B.: Existence and nonuniqueness of local separation zones in viscous jets, Izv. Akad. Nauk SSSR: Mekh. Zhidk. Gaza 21, 1 (1986) 38-45.

Address: PROF. DR. ALFRED KLUWICK, Vienna University of Technology, Institute of Fluid Dynamics and Heat Transfer, Wiedner Hauptstraße 7, A-1040 Vienna, Austria.



DANIEL MARGERIT AND DWIGHT BARKLEY

## Singular perturbation equations for 3-d excitable media

*In this paper the idea of Prandtl's boundary layer is exported to a field other than fluid dynamics. Excitable media, such as nerve fibers and heart tissue, are typically modelled with reaction-diffusion equations containing two chemical species that evolve on very different time scales. In three dimensions solutions of these equations take the form of rotating scroll waves (interfaces) ending on filaments. The ratio of the two times scale defines a natural small parameter epsilon. Exploiting the inherent smallness of epsilon, singular perturbation methods are used to derive three-dimensional equations for each of two boundary layers : interface region (scroll) and filament region (core), and for the associated outer region. For scrolls with uniform twist about straight filaments, this matched asymptotic expansion method is also used to derive free-boundary equations not only at leading order but also at first order. Both orders are validated against full solutions of the reaction-diffusion equations. Using these two orders and with no adjustable parameters, the shape and frequency of waves are correctly predicted for most cases of physical interest.*

### 1. Introduction

In three-dimensional excitable media, propagating waves of excitation typically take the form of scrolls which are organized about one-dimensional filaments[13]. These filaments have some similarities to the vortex filaments found in fluid dynamics. However, unlike vortex filaments in fluid dynamics, filaments in excitable media can have associated *twist*. Figure 1 illustrates this by showing a scroll wave which is uniformly twisted along a straight filament. The purpose of this paper is to show how the idea of Prandtl's boundary layer can be applied to excitable media. In particular we derive equations predicting the shape and rotation frequency of scroll waves such as in Fig. 1 and through these equations we are able to understand and predict the role of twist in shape and frequency selection.

We begin by considering the following partial-differential-equation (PDE) model of excitable media[1] written in the space-time scales proposed by Fife[6]:

$$\epsilon^2 \partial u / \partial t = \epsilon^2 \nabla^2 u + u(1-u) \left( u - \frac{v+b}{a} \right), \quad (1)$$

$$\partial v / \partial t = \epsilon(u-v). \quad (2)$$

Such two-component reaction-diffusion models capture essential properties of excitable media and are widely used in theoretical and computational studies, e.g. [2, 3, 7, 8, 9, 14]. Model parameters  $a$  and  $b$  control the excitation threshold and duration and will have values  $a = 0.8$  and  $b = 0.1$  throughout. The parameter  $\epsilon$  is small, reflecting the disparate time scales of the fast activator variable  $u$  and slow inhibitor variable  $v$ .

Previous work on wave selection in excitable media through asymptotic expansions [3, 7, 8, 9, 12] has focused entirely on leading order in the small parameter  $\epsilon$  and primarily on two dimensions. Expanding the rotation frequency as

$$\omega = \omega^{(0)} + \epsilon \omega^{(1)} + \dots, \quad (3)$$

only the leading-order frequency  $\omega^{(0)}$  has been obtained[3, 7]. While the small- $\epsilon$  (Fife) limit has played an important role in 2D studies, the leading order does not accurately predict many properties of waves at finite  $\epsilon$ . However, we find (Fig. 3 below) that expansions to *first order* in  $\epsilon$  are predictive well into regimes of physical interest.

### 2. Geometry, asymptotic description and leading order solution

For the leading-order asymptotics, we begin by considering the general three-dimensional (3D) case. The medium is divided into three regions: outer, interface, and core as shown in Fig. 2. The filament is the curve  $\mathbf{X}(s, t)$  inside the core. The outer region comprises the bulk of the medium. It consists of both excited (+) and quiescent (-) portions for which  $u = u^+ = 1$  and  $u = u^- = 0$ , respectively, to all orders in  $\epsilon$ . Expansion of the  $v$ -field in the outer region gives:  $v = v^s + \epsilon v^{(1)} + \dots$ , where  $v^s = -b + a/2$  is the stall concentration (value such that a plane interface is stationary) and  $v^{(1)}$  is to be determined.

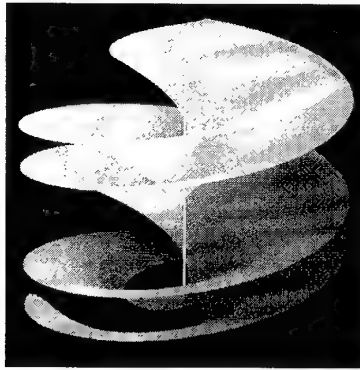


Figure 1: Twisted scroll wave from numerical solutions of Eqs. (1-2). Isosurface is shown for  $u = 0.5$ . The filament is white. The structure rotates in time with frequency  $\omega$  about the filament. The twist is  $\tilde{\tau} = 0.5$ , (defined later in the text);  $\epsilon = 0.1$ .

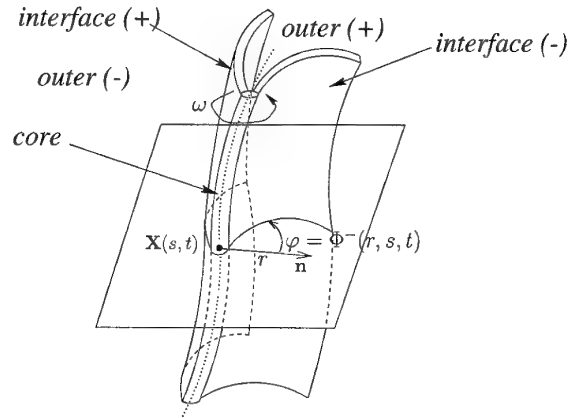


Figure 2: Scroll geometry showing outer regions [excited (+) and quiescent (-)], interface regions [wavefront (+) and waveback (-)], and core region. The filament  $\mathbf{X}(s, t)$  is parameterized by  $s$  and time  $t$ . Local coordinates to the filament are  $(r, \varphi, s)$ , with  $(r, \varphi)$  in the plane normal to  $\mathbf{X}(s, t)$  and  $\varphi$  measured from the normal vector  $\mathbf{n}$ .

Separating excited and quiescent states are the thin interfaces where  $u$  undergoes rapid change. These consist of a wave front (+) and a wave back (-), which on the outer scale are given by  $\varphi = \Phi^\pm(r, s, t)$ . Solving leading- and first-order inner equations for  $u$  across the interface ( $v$  is constant at these orders across the interface) and matching to the outer  $u$ -solution, one obtains equations for interface motion[10]. Thus Eqs. (1-2) reduce to equations for  $v^{(1)}$  in the outer region together with equations for the motion of the two interfaces (free boundaries):

$$\partial v^{(1)}/\partial t = u^\pm - v^s, \quad (4)$$

$$-\frac{r\dot{\Phi}^{(0)\pm}h^\pm}{\sqrt{m^\pm}} = 2H^\pm \pm \frac{\sqrt{2}}{a}v^{(1)\pm} \quad (5)$$

where  $\Phi^{(0)\pm}$  is the leading order approximation to  $\Phi^\pm$ , and where  $h^\pm \equiv |\partial\mathbf{X}/\partial s|(1 - rK \cos \Phi^{(0)\pm})$ ,  $K$  is the filament curvature,  $m^\pm$  is the determinant of the metric tensor and  $H^\pm$  is the mean curvature of interface  $\Phi^{(0)\pm}$ , and finally  $v^{(1)\pm}$  is the value of  $v^{(1)}$  at interface  $\Phi^{(0)\pm}$ . Eq. (5) equates normal velocity of the interface to twice the mean curvature plus the speed of a plane interface. Phenomenological approaches to excitable media yield similar equations [16]. As in 2D [9], the core plays no role at leading order other than to regularize the cusp that would otherwise exist as the two interfaces come together. However, leading-order core equations dictate that  $\dot{\mathbf{X}}^{(0)} = 0$ , i.e. the filament velocity is zero at leading order in  $\epsilon$  and filament motion must come at higher order.

We now consider the specific case of a straight filament and seek solutions with uniform twist  $\tau \equiv \partial\Phi/\partial s$  and constant frequency  $\omega^{(0)} = \dot{\Phi}^{(0)}$ . The angle between the two interfaces can be shown to be constant:  $\Delta\Phi^{(0)} = \Phi^{(0)-} - \Phi^{(0)+} = 2\pi(1 - v^s)$  and  $v^{(1)\pm}$  can be eliminated from the free-boundary equations to obtain a single equation describing the shape of the interface[10]:

$$q \frac{d\Psi^{(0)}}{d\tilde{r}} + \frac{\Psi^{(0)}(1 + \Psi^{(0)2})}{\tilde{r}} = \tilde{r}(q + \Psi^{(0)2}) - B(q + \Psi^{(0)2})^{3/2}, \quad (6)$$

where  $\Psi^{(0)} \equiv rd\Phi^{(0)}/dr = rd\Phi^{(0)}/dr$ , and  $q \equiv 1 + \tilde{\tau}^2\tilde{r}^2$ , with  $\tilde{r} = \sqrt{\omega^{(0)}}r$ ,  $\tilde{\tau} = \tau/\sqrt{\omega^{(0)}}$ . The eigenvalue  $B$  is related to  $\omega^{(0)}$  and model parameters via  $B = (\mu/\omega^{(0)})^{3/2}$  where  $\mu^{3/2} = \sqrt{2}\pi v^s(1 - v^s)/a$ . With  $\tilde{\tau} = 0$  (2D case), Eq. (6) is as given by Karma[7], while for  $\tilde{\tau} \neq 0$  it can be shown to agree with the work of Bernoff [3].  $\Psi^{(0)}$  and the selected  $B$  as a function of  $\tilde{\tau}^2$  is found[11] from Eq. (6) by shooting: integrating from  $\tilde{r} = 0$  to large  $\tilde{r}$  and finding  $B$  such that  $\Psi^{(0)}$  matches the relevant large- $r$  limit obtained from Eq. (6).

### 3. Order- $\epsilon$ asymptotic for a straight filament with twist

We now consider the order- $\epsilon$  asymptotics. We treat only the case of straight filaments. For scrolls with twist  $\tau$  rotating at frequency  $\omega$ :  $\partial/\partial t = -\omega\partial/\partial\varphi$  and  $\partial/\partial z = \tau\partial/\partial\varphi$ . For this case Eqs. (1-2) become

$$\epsilon^2\omega\partial u/\partial\varphi + \epsilon^2\nabla_{\perp}^2 u + u(1-u)\left(u - \frac{v+b}{a}\right) = 0, \quad (7)$$

$$\omega\partial v/\partial\varphi + \epsilon(u-v) = 0, \quad (8)$$

where  $\nabla_{\perp}^2 = \partial^2/\partial r^2 + (1/r)\partial/\partial r + (q/r^2)\partial^2/\partial\varphi^2$ . The major complication in deriving free-boundary equations from Eqs. (7-8) is matching outer and inner solutions (for  $u$  and  $v$ ) across the interface because  $v^{(2)}$  is not constant across the interface and because the normal to the interface lies outside the  $(r, \varphi)$  plane when  $\tau \neq 0$ . For this we use local coordinates "normal" to  $\nabla_{\perp}^2$  near the interface[10].

The symbolic calculator Maple is used to obtain the cascade of asymptotic equations in both the outer and inner regions up to the order of interest. The outer asymptotic expansion is plugged into Eqs. (7-8). For the inner region, Maple is first used to express Eqs. (7-8) in local inner coordinates in the interface region and then used to plug the inner asymptotic expansions into these equations. Maple is then used to find the behavior at infinity of the inner solution and to perform the intricate asymptotic matching with the outer solution. Finally, Maple is used to find the asymptotic behavior at infinity of  $\Psi^{(0)}$  and  $\Psi^{(1)}$  to many orders in  $\tilde{r}$ . The symbolic calculator allows us to quickly derive these results and to minimise the possibility of mistakes in such fastidious calculus.

The result is that at this order  $\Delta\Phi^{(1)} = \Phi^{(1)-} - \Phi^{(1)+} = 0$  and it is again possible to obtain a single equation for  $\Psi^{(1)} \equiv a\omega^{(0)}r d\Phi^{(1)\pm}/dr$ [11] with an eigenvalue  $D$  related to  $\omega^{(1)}$  by  $D = a\omega^{(1)}$ . The general solution of this equation is found[11] and diverge exponentially at infinity unless  $D$  has a selected value.

Finally, we use symbolic calculation to verify that the fields obtained asymptotically is truly the solution of Eqs. (7-8) up to the order of interest. This is an exact check which is independent of the calculus used to derived the asymptotic fields. We perform this verification with Maple by (i) plugging into Eqs. (7-8) the outer asymptotic solution and equations for  $\Psi^{(0)}$  and  $\Psi^{(1)}$  up to the relevant order, then expanding in  $\epsilon$  and verifying that  $0 = 0$  on the computer; (ii) doing the same for the inner asymptotic solution and with Eqs. (7-8) written in the local stretched coordinates (but not expanded in  $\epsilon$ ); (iii) checking the matching between the outer and inner asymptotic solutions. Such a check is important in boundary layer problems. The symbolic calculator makes this check easy to perform and thus provides a strong and useful tool for singular perturbation calculus.

### 4. Comparison with the numerical PDE solution and conclusion

We now compare the asymptotic results with full PDE solutions. For this we solve (7-8) using Newton's method [2]. The operator  $\nabla_{\perp}^2$  is discretized on a polar grid typically with 256 points in  $\varphi$  and radial spacing  $\Delta r = 0.05$ . The  $r$ -derivatives are computed by finite differences and  $\varphi$ -derivatives are computed spectrally.

Figure 3 shows the dependence of  $\omega$  on  $\epsilon$  from the PDE solutions. This figure clearly shows the existence of the Fife limit: a finite-frequency limit as  $\epsilon \rightarrow 0$ . Over a substantial range of  $\epsilon$ , the frequency is very well captured by the first two orders in  $\epsilon$ :  $\omega \simeq \omega^{(0)} + \epsilon\omega^{(1)}$ . Extrapolation of frequency data to  $\epsilon = 0$  gives  $\omega^{(0)}$  and thus  $B$ . The slope of  $\omega$  versus  $\epsilon$  gives  $\omega^{(1)}$  and hence  $D$ . From the computed  $u$ -fields we find the functions  $\Phi^{\pm}$  as curves on which  $u = 1/2$  and from these  $\Psi$  is computed by differencing. Analogously to the frequency, from the dependence of  $\Psi$  on  $\epsilon$  we find  $\Psi^{(0)}$  and  $\Psi^{(1)}$ [11]. The core radius is found to be  $r \simeq 8\epsilon$  and the data also confirm that  $\Delta\Phi^{(1)} = 0$ .

In Fig. 4 we compare full solution of the stationary PDE (7-8) with the interface curves. Shown is a cross-section of a twisted scroll wave normal to the straight filament  $\mathbf{X}$  at station  $s$  and instant  $t$  in the domain,  $r \leq 20$ . Also shown is cross-section of the stationary interface at leading-order  $\varphi = \Phi^{(0)\pm}(r, s, t)$  and at leading-plus-first-order  $\varphi = \Phi^{(0)\pm}(r, s, t) + \epsilon\Phi^{(1)\pm}(r, s, t)$ . Figures 4(b) and (d) show the same case as Fig. 1 (apart from the domain radius). The agreement is excellent and contains no adjustable parameters.

In conclusion, we have derived free-boundary equations at leading-order and first-order for twisted scroll waves in excitable media and we have validated these equations directly with numerical solutions of the underlying PDEs. The free-boundary equations we have derived apply to a large class of models [10]. For excitable media it would be of considerable interest to derive an equation of motion for this filament as has been successfully performed in hydrodynamics for vortex filaments[4].

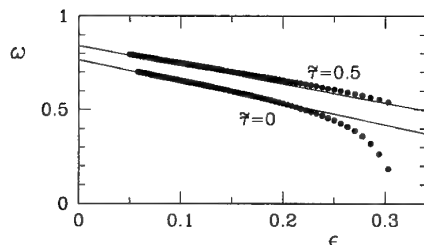


Figure 3: Scroll frequency  $\omega$  versus  $\epsilon$  from numerical solutions of the PDE model for two values of twist. Lines are from fits to the data at small  $\epsilon$  and are indistinguishable from asymptotic predictions at first order in  $\epsilon$ .

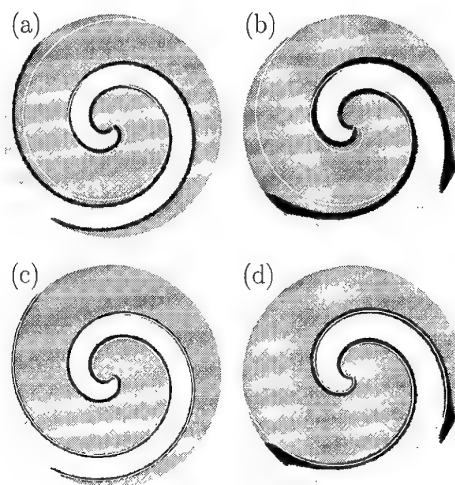


Figure 4: Comparison between PDE solutions (greyscale) and asymptotic results (white curves). (a)  $\tilde{\tau} = 0$ , asymptotics at leading order. (b)  $\tilde{\tau} = 0.5$ , asymptotics at leading order. (c)  $\tilde{\tau} = 0$ , asymptotics at leading-plus-first order. (d)  $\tilde{\tau} = 0.5$ , asymptotics at leading-plus-first order. Black is the interface  $0.1 \leq u \leq 0.9$ ; light grey (dark grey) is  $u < 0.1$  ( $u > 0.9$ ). The radius is 20;  $a = 0.8$ ,  $b = 0.1$ ,  $\epsilon = 0.1$ .

## 5. References

- 1 BARKLEY, D.: A model for fast computer simulation of waves in excitable media; *Physica D* **49** (1991), 61–70.
- 2 BARKLEY, D.: Linear Stability Analysis of Rotating Spiral Waves in Excitable Media; *Physical Review Letters* **68** No. 13 (1992), 2090–2093.
- 3 BERNOFF, A.J.: Spiral wave solutions for reaction-diffusion equations in a fast reaction/slow diffusion limit; *Physica D* **53**(1991), 125–150.
- 4 CALLEGARI, A.J., TING, L.: Motion of a curved vortex filament with decaying vortical core and axial velocity; *SIAM J. Appl. Math.* **35** No. 1, (1978), 148–175.
- 5 FENTON, F., KARMA, A.: Vortex dynamics in 3 dimensional continuous myocardium with fiber rotation : filament instability and fibrillation; *Chaos* **8** No. 1 (1998), 20–47.
- 6 FIFE, P.C.: Understanding the Patterns in the BZ Reagent; *Journ. of Statistical Physics* **39** No. 5/6 (1985), 687–703.
- 7 KARMA, A.: Scaling Regime of Spiral Wave Propagation in Single-Diffusive Media; *Physical Review Letters* **68** No. 3 (1992), 397–400.
- 8 KEENER, J.P., TYSON, J.J.: The Dynamics of Scroll Waves in Excitable Media; *SIAM Review* **34** No. 1 (1992), 1–39.
- 9 KESSLER, D.A., LEVINE, H., REYNOLDS, W.: Theory of the spiral core in excitable media; *Physica D* **70** (1994), 115–139.
- 10 MARGERIT, D., BARKLEY, D.: to be published (2000).
- 11 MARGERIT, D., BARKLEY, D.: Selection of twisted scroll waves in three-dimensional excitable media; to appear in *Phys. Review Letters* (2000).
- 12 MIKHAILOV, A.S., ZYKOV, V.S.: Kinematical theory of spiral waves in excitable media: Comparison with numerical simulations; *Physica D* **52** (1991), 379–397.
- 13 WINFREE, A.T.: Stable Particle-Like Solutions to the Nonlinear Wave Equations of Three-Dimensional Excitable Media; *SIAM Review* **32** No. 1 (1990), 1–53.
- 14 WINFREE, A.T.: Varieties of spiral wave behavior: An experimentalist's approach to the theory of excitable media; *Chaos* **1** (1991), 303–334.
- 15 WINFREE, A.T.: Persistent tangles of vortex rings in excitable media; *Physica D* **84** No. 1–2 (1995), 126–147.
- 16 YAMADA, H., NOZAKI, K.: Dynamics of untwisted scroll rings in excitable media; *Journal of the physical society of Japan* **63** No. 2 (1994), 379–382.

Addresses: DR. DANIEL MARGERIT, DR. DWIGHT BARKLEY, Mathematics Institute, University of Warwick, Coventry CV4 7AL, UK

TING, LU AND VAN DYKE, MILTON

## Prandtl's Boundary Layer Theory, before and to Matched Asymptotics

*The idea of joining a local and a global perturbation expansion to form a uniformly valid approximation was exploited by various workers in mechanics throughout the 19th century for isolated problems. The idea was first explored systematically by Ludwig Prandtl with his students and colleagues, for fluid motion at high Reynolds number. Here we review the work of Prandtl's 19th-century predecessors and then describe the explanation of Prandtl's boundary layer theory by K. O. Friedrichs, leading to the method of matched asymptotics. The method provides a systematic procedure for the formulation of a singular perturbation problem, and additional mathematical novelties, such as, the identification of the 'lost' boundary condition(s) from the full system of equations and the derivation of the next order equations and their compatibility condition(s), which in turn serve as the closure condition(s) for the leading order solution. Examples are presented to demonstrate the novelties of matched asymptotics and to emphasize the physical intuition needed to formulate the perturbation problem, i. e., the choice of the scalings and the expansion schemes.*

### 1 Introduction

In 1904 [1], Prandtl presented his boundary layer theory for a body moving at high Reynolds number. See also [2]. His theory initiated a systematic procedure for joining local (inner) and global (outer) perturbation expansions to form a uniformly valid approximation, and was generalized by Friedrichs in 1945, [3], to a systematic procedure for deriving the leading and higher order equations and matching conditions for the inner and outer solutions and for removing singularities of the outer solutions. The procedure was referred to as the boundary layer technique [4], and became known since the 60's as the method of matched asymptotics (MMA). The method has been widely employed to resolve many singular perturbation problems in applied mechanics and have been explained in many monographs, see, e. g., [5] and [6].

The first part of our presentation in the Mini-symposium, dealing with the boundary layer type analyses prior to Prandtl, is described in Section 2. We mention problems in fluid dynamics and acoustics. for which the construction of a global or regular perturbation solution with local singularities was carried out, and the solution was later identified as the leading global solution by MMA, with its singularities resolved by the inner solutions.

The second part of our presentation reviews the formulation of and the contributions to MMA by Prandtl and Friedrichs and their students and colleagues. The review emphasizes the physical intuition needed to formulate the perturbation problem, i. e., setting up the expansion scheme, the restrictions implied by the expansion scheme and the physical meaning of the inner solution and its matching with the outer solution. The highlights of the review article, [7], are presented in Section 3.

The last part of our presentation is described in Section, 4. We use the studies of the diffraction of weak shock by a concave corner to show that there can be different admissible models, in the sense that the matched asymptotic analysis for each model can be carried out to higher orders without contradiction. But different models have different degree of difficulty, sometimes insurmountable, in the construction of the inner solution. We then identify the "best" model for the removal of a particular singularity.

### 2 Boundary Layer Type Analyses Before Prandtl

We note that the analysis of a linear oscillator with vanishing mass was the example introduced by Prandtl to explain his boundary layer theory [2] and was explained in detail in [6] to show the basis of MMA. We shall not

repeat this example. Instead, we mention one classical problem in fluid dynamics, and one in acoustics in §2.1 and 2.2 respectively and interpret them by MMA.

**Point vortex in a two-dimensional potential flow.** The classical solution for the velocity potential a point vortex in a background flow is  $\Phi(x, y, t) = [\Gamma/2\pi] \ln r + \phi(x, y)$ , where  $\phi(x, y)$  is the potential of the steady background flow without the vortex, and  $\Gamma$  denotes the strength of the vortex located at point  $P(X(t), Y(t))$  and  $r$  denotes the distance from  $(x, y)$  to  $P$ . This solution has two defects: (1) the solution is singular when  $r \rightarrow 0$  and (2) the velocity of the vortex point has to be assigned. It is zero for a fixed vortex and is the local background,  $\dot{P}(t) = \nabla\phi(X, Y)$  for a free vortex. The solution for a fixed vortex is often identified as the far field representation of a lifting body of size,  $b$ . The inner solution, or the near field solution in the length scale,  $b$ , yields or matches with the outer solution at large distance  $L$  to the body with  $L \gg b$  [8].

When there is a free vorticity distribution of total strength  $\Gamma$ , concentrated in the neighborhood of a point  $P$ , the classical solution is considered to be the outer solution in the length scale  $\ell$  much larger than the effective core size  $\delta$  of the vortex distribution. The singularity of the outer solution as  $r/\ell \rightarrow 0$  is removed or matched with the inner solution, i. e., the core structure. With the core structure, the velocity of the vortex is defined and the classical theory for the velocity of the point vortex is identified by MMA as the average velocity in the normal time scale  $O(\ell^2/\Gamma)$ . See [9] and references therein.

**2.2. Scattering of long acoustic waves** The problems of acoustic radiation from the open end of a pipe, apertures in plane screens and scattering by small obstacles were treated by Helmholtz (1860) and Rayleigh (1897). See [8]. The solution, known as Rayleigh's long wave approximation, is applicable when the wave length  $\lambda$  is much larger than the size  $a$  of the opening or the obstacle. The scattered field in the length scale  $a$ , obeys the Laplace equation with  $t$  as a parameter, instead of the wave equation. Rayleigh's approximation is explained systematically by MMA with  $a/\lambda$  as the expansion parameter and applied to other problems in acoustics by Lesser and Lewis [10] and Ting and Keller [11], and to problems in aero-acoustics by Crow (1970) and many others (see [8] and the references therein).

### 3 Boundary Layer Theory to Matched Asymptotics

In the review paper [7], we describe the formulation of the boundary layer theory in four steps, common to those used later in MMA, [3] and [5], in which the boundary layer solution and the inviscid solution are identified as the leading order inner and outer solutions respectively. The method gives a systematic procedure to derive the higher order equations for the inner and outer solutions and their matching conditions. The matching conditions also resolve the local singularities of the outer solution, if any. The four steps are: *I*, the physical intuition or modeling of the flow field, *II*, the choice of the scalings and the expansion scheme, *III*, derivation of the leading and higher order equations and the matching conditions, and *IV*, the construction of the inner and outer solutions and the study of their physical meaning. In terms of these four steps, the method enables us to check the consistency of the expansion scheme and provides additional mathematical novelties, such as the "lost" boundary conditions, the "compatibility" conditions and the "closure" conditions.

We point out that the selection of an the expansion scheme in step *II* imposes certain restrictions on the solution. Two examples are quoted to show the modifications needed when some of the restrictions are violated. The consistency of an expansion scheme has to be tested in step *III*, not only by the leading and higher order equations of the inner and outer solutions but also by their matching conditions. For example, in Kaplun's analysis of a flow at low Reynolds number,  $Re \ll 1$ , the necessity of a double series expansion in power of  $Re$  and  $\ln(1/Re)$  comes from the matching conditions. See [5]. From step *III*, we can identify the "lost" boundary condition(s) from the full system of equations and the associated compatibility condition(s) for the solutions of the reduced systems. As examples, we identify the "lost" boundary conditions and the "compatibility" conditions for two-dimensional and three-dimensional boundary layers. We mention the second order shock conditions to demonstrate that there are cases for which we have to recover the missing or "closure" conditions for the leading order solutions from compatibility conditions of the higher order equations.

We note that although Prandtl did not continue his boundary layer theory to the next order, he introduced the displacement and momentum thicknesses of the boundary layer and pointed out how to use them to render the outer solution valid to the next order, [12]. His suggestions have been in use by engineers, for example, the design

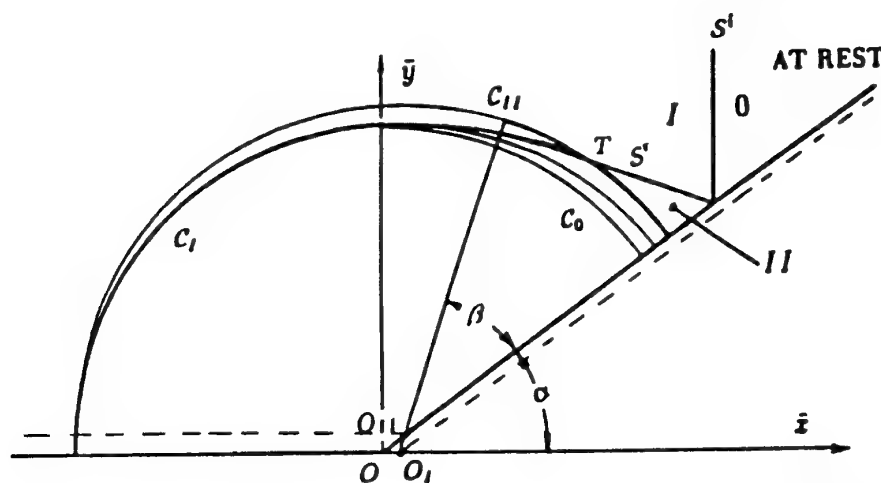


Figure 1: Shock incident on a concave corner - uniform regions, sonic circles and shock fronts.

of convergent-divergent nozzles for supersonic tunnels and the design of supercritical wings for current commercial airplanes.

We then describe the investigation leading to the correct model for the planing of a plate at high Froude number in order to show that it is essential to understand the physics of the problem prior to the formulation of the correct model and expansion scheme. In the last section of [7], we mention a few methods for singular perturbation problems, other than MMA and note that MMA can be readily implemented to remove the singularities of the solutions obtained by the other methods.

## 4 Diffraction of Weak Shocks

We use the studies of the diffraction of a weak shock by a concave corner to demonstrate that there can be different models for MMA and identify the "best" model for the removal of a particular singularity of the outer solution. Let the negative  $x$ -axis be the horizontal side of a concave corner of angle  $\pi - \alpha$ , with vertex at the origin  $O$  and the inclined side be parallel to  $\hat{r} = \hat{i} \cos \alpha + \hat{j} \sin \alpha$ . The incident shock is moving parallel to the  $x$ -axis with velocity  $V_0 \hat{i}$  hitting the inclined side at  $t = 0$ . For  $t < 0$ , there is no reflected shock. For  $t > 0$ ,  $S^i$  is reflected by the inclined side at the point  $V_0 t \sec \alpha \hat{r}$  and  $\beta$  denotes the angle of reflection of the shock  $S^r$ . There are three uniform regions, 0, I and II, outside the non-uniform region in the domain of influence of the vertex. Ahead of  $S^i$ , there is the uniform region 0, at rest. Behind  $S^i$  but ahead of the reflected shock  $S^r$ , there is the uniform region I, moving at velocity  $u_{II} \hat{i}$ . Behind  $S^r$ , there is the uniform region II, moving at velocity  $u_{II} \hat{r}$ . Let  $P_k$  and  $C_k$  denote the pressure and speed of sound of the  $k$ -th uniform state, in which the domain of influence of the vertex is bounded by the sonic circle  $C_k$  with radius  $C_k t$  and center  $O_k$  moving with the uniform stream. We have  $O_0$  located at the origin,  $O_I$  at  $iu_{II}t$  and  $O_{II}$  at  $\hat{r}u_{II}t$ . We note that, (i) relative to a shock, the flow behind (ahead) or on the high (low) pressure side is subsonic (supersonic) and hence (ii)  $S^r$  cuts into  $C_{II}$  at point  $T$ , as a diffracted shock  $S^d$ , approaching  $C_I$  from outside and remains outside of  $C_0$ . Since the unsteady flow field does not have a length scale, it is a conical flow field. Figure 1 shows the flow field in the conical variables  $\bar{x} = x/C_0 t$  and  $\bar{y} = y/C_0 t$ . The uniform regions, 0 and II, being outside of the domain of influence of the vertex, are defined. The uniform region I is behind  $S^r$ , above the  $x$ -axis and outside the reflected shock  $S^r$  and then the diffracted shock,  $S^d$ , yet to be defined.

For a weak incident shock, we have  $(P_I - P_0)/P_0 = \epsilon \ll 1$ , and  $\beta = \alpha + O(\epsilon)$ , and the flow field differs from any one of the three uniform states by  $O(\epsilon)$ . The classical theory for weak shocks or acoustic wave gives the perturbation solution from the state at rest, say the perturbation pressure  $\epsilon p = P - P_0$ , given by Keller and Blank (1951), cited in [13]. In the conical variables, the non-uniform region is bounded by the unit circle,  $C_0$ , and the sides of corner,  $\theta = \pi$  and  $\pi - \alpha$ . In the first approximation, the reflected shock  $S^r$  is tangent to  $S_0$  at  $T(1, 2\alpha)$ , which is the triple point lying on the boundaries I and II and the non-uniform region. Across the point  $T$  along  $C_0$ , the pressure jumps from  $P_I$  to  $P_{II}$  as  $\theta$  decreases. We note that, (iii) in conical variables, the linearized equation is hyperbolic outside and elliptic inside the unit circle.



The perturbation solution is not uniformly valid to  $O(\epsilon)$  near  $C_0$  because it differs from correct boundaries of its adjacent uniform regions,  $C_I$  and  $C_{II}$ , by  $O(\epsilon)$  and  $\nabla p$  is singular near  $C_0$ . The solution was rendered uniformly valid excluding the neighborhood of the triple point by Lighthill's technique (1949) and by MMA, Zahalak and Myers (1974), for linearized supersonic flows and by the geometrical theory of diffraction for nonlinear waves, Hunter and Keller (1987). See [13]. To  $O(\epsilon)$ , they arrived at the same correction which amounts to replacing the arcs of  $C_0$ , adjacent to  $I$  and  $II$  and the distance to the arcs by those of  $C_I$  and  $C_2$  respectively.

If we construct the perturbation solution say  $\epsilon p_k = P - P_k$  from the uniform state,  $k = I$  or  $II$ , and use  $C_k$  as the origin for the conical variables  $(\bar{x}_k, \bar{y}_k)$ , we find that  $\epsilon p_k = \epsilon p(\bar{x}_k, \bar{y}_k) + (P_0 - P_k)$ , provided that the vertex  $C_0$  is moved to  $C_k$ , i. e., the horizontal (inclined) side is moved with the uniform stream in  $I$  ( $II$ ) to the dotted line as shown in Fig. 1. Thus  $p_k$  is related to  $p$  in the  $k$ -th conical variables. We note that  $p_I$  ( $p_{II}$ ) is uniformly valid near  $C_I$  ( $C_{II}$ ). In the region  $\mathcal{R}$  common to the domains of the three perturbation solutions, which is outside the  $\epsilon$ -neighborhood of the sonic circles and the sides of the corners, we have  $P_k + \epsilon p_k$ ,  $k = I, II$  and  $P_0 + \epsilon p$  differ by  $O(\epsilon^2)$ , because their conical variables differ by  $O(\epsilon)$ . Excluding the neighborhood of the triple point, we have a perturbation solution uniformly valid to  $O(\epsilon)$ , which is  $p(\bar{x}_k, \bar{y}_k)$  near  $C_k$  for  $k = I$  or  $II$  and continue to  $p(\bar{x}, \bar{y})$  in  $\mathcal{R}$  and then to the corner.

Near the triple point  $T$  or the singular ray, the correction, breaks down because of the  $O(\epsilon)$  gap between  $C_I$  and  $C_{II}$  and pressure jump from  $P_I$  to  $P_{II}$ . The inner solution for the neighborhood of a singular ray,  $\mathcal{N}$ , was found to obey a transonic equation by the aforementioned three methods. Because of statements (ii) and (iii), the solution perturbed from state 0 is of the mixed type and is not available. This is also true for that perturbed from state  $I$ . For the solution perturbed from state  $II$ , the inner solution in  $\mathcal{N}$  remains elliptic and its solution is amenable [14]. From the physics of shock waves and characteristics, we arrive at the following rule: *At a singular ray separating two uniform states, the perturbation expansion based on the state with the higher pressure is the "best" one, because it leads to an elliptic problem in  $\mathcal{N}$ .*

### Acknowledgements

The research of L. Ting was partially supported by the Alexander von Humboldt Foundation.

### References

- 1 PRANDTL, W.: Über Flügigkeitsbewegung bei sehr kleiner Reibung; In *Verh. int. Math. Kongr., Heidelberg*, pages 481-491. Teubner, Leipzig, 1904.
- 2 SCHLICHTING, H. AND GERSTEN, G.: *Boundary Layer Theory*; 8th Ed., Springer Verlag, NY, 1999.
- 3 VON MISES, R., AND FRIEDRICHS, K. O.: *Fluid Dynamics*; Brown University Lecture Notes, 1946. Reprinted as *Applied Mathematical Sciences*, 5, Springer-Verlag, New York, 1971.
- 4 FRIEDRICHS, K. O.: Asymptotic phenomena in mathematical physics; *Bull. Amer. Math. Soc.* **61**, (1955), 485-504.
- 5 VAN DYKE, M.: *Perturbation Methods in Fluid Mechanics*; The Parabolic Press, Stanford, 1975.
- 6 KEVORKIAN, J. AND COLE, J. D.: *Perturbation Methods in Applied Mathematics*; Springer Verlag, NY, 1980.
- 7 TING, L.: Boundary layer theory to matched asymptotics; to appear in the special issue of ZAMM dedicated to Ludwig Prandtl, 2000.
- 8 LAMB, H.: *Hydrodynamics*, Cambridge University Press, 1st. ed. (1879), 6th ed. (1932).
- 9 TING, L. AND KLEIN, R.: Viscous Vortical Flows; *Lecture Notes in Physics*, **374** Springer-Verlag, NY, 1991.
- 10 LESSER, M. B. AND LEWIS, J. A.: Applications of matched expansion methods to acoustics, I, The Webster equation and the stepped duct, and II, The open ended duct; *J. Acoust. Soc. Am.*, **51**, (1972), 1664-1669 and **52**, (1972), 1406-1410.
- 11 TING, L. AND KELLER, J. B.: Radiation from the openend of a cylindrical or conical pipe and scattering from the end of a rod or slab; *J. Acoust. Soc. Am.*, **61**, (1977), 1438-1444.
- 12 PRANDTL, L. AND TIETJENS, O. G.: *Applied Hydro- & Aeromechanics*; Dover Publ., NY, 1934.
- 13 TING, L. AND KELLER, J. B.: Weak shock diffraction and singular rays; *ZAMM* **78**, (1998), S767-770.
- 14 TING, L. AND KELLER, J. B.: Weak diffracted shocks near singular rays; to appear in *J. Methods and Applications of Analysis*, 2000.

Address: PROF. LU TING, New York University, Courant Inst., 251 Mercer Street, New York, NY 10012, USA.  
 PROF. MILTON VAN DYKE, Stanford University, Div. of Mech., Stanford, CA 94305, USA.



MACHT, J.; LACKNER, R.; HELLMICH, CH.; MANG, H.A.

## Geomechanics of tunneling in squeezing rock

*This paper deals with the numerical analysis of shotcrete tunnel shells in squeezing rock. For this purpose, a hybrid method combining in-situ measurements with chemomechanical material modeling of shotcrete is developed.*

### 1. Tunneling in Squeezing Rock

When driving tunnels in squeezing rock, the ground squeezes (plastically) into the opening without visible fracturing or loss of continuity [1]. In the context of the New Austrian Tunneling Method (NATM), the newly excavated tunnel area is supported by means of a thin flexible shell of shotcrete. Under "normal" ground conditions, the compliance of the shotcrete shell resulting from elastic, plastic, and creep deformations is sufficient to cope with the moderate ground movements. For squeezing rock, however, the compliance of the shell is not sufficient to avoid damage or destruction of the shell. In order to increase the compliance of the shotcrete shell, longitudinal gaps are left out during shotcreting. In these gaps, so-called Lining Stress Controllers (LSC) may be placed. Figure 1(a) shows one of three rows of LSCs installed at the Semmering pilot tunnel. LSCs are used to control the load level of the shotcrete shell. At a pre-specified load level, buckling occurs which is followed by local plastic deformations (see Figure 1(b)). The experimentally obtained load-displacement curve for the LSCs used at the Semmering pilot tunnel is shown in Figure 1(c).

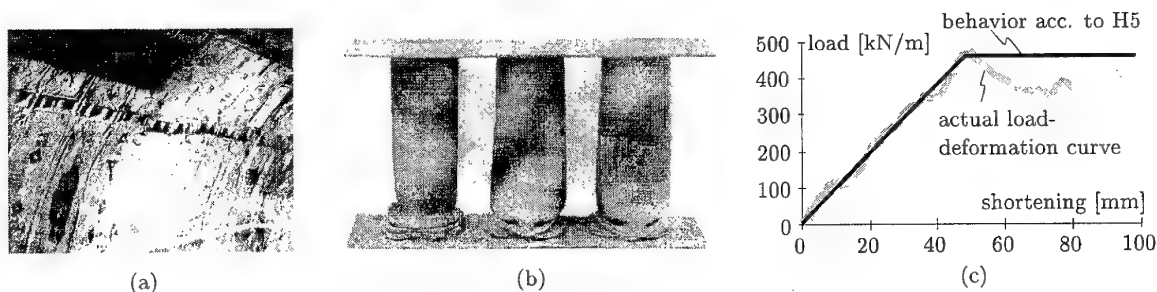


Figure 1: Semmering pilot tunnel: (a) shotcrete tunnel shell with Lining Stress Controllers (LSC) installed in longitudinal gaps [6], (b) prototype of LSC after buckling [6], and (c) experimentally obtained load-displacement curve for LSCs used at Semmering pilot tunnel [6]

### 2. Hybrid Method for Tunneling in Squeezing Rock

Hybrid methods have been developed in order to gain realistic informations concerning the state of loading of tunnel shells [5] [3]. These methods are characterized by prescribing displacements measured *in situ* on the exterior surface of the structural model of the investigated part of the tunnel shell. The investigation concerns a ring with a width of 1 m, fictitiously cut out of the structure and modelled, in the present case, by plane-strain Finite Elements. The transfer of the measured displacements to the structural model requires a couple of hypotheses concerning the structural behavior of a tunnel shell. As for a shotcrete shell with LSCs, the following hypotheses are made:

- H1: During the deformation, the *thickness* of the shell is assumed to be approximately *constant*. This hypothesis is consistent with the Kirchhoff-Love shell theory. Because of the small thickness and the mode of loading of the shell, this assumption is justified.
- H2: As for the parts of the shell made of shotcrete, *smooth displacement fields* are assumed. With regards to the gaps, after buckling of the LSCs (see localized deformation in Figure 1(b)), large displacement gradients occur in a very small region as compared to the dimensions of the whole structure. This situation can be approximated by *discontinuous* displacement fields, i.e., they exhibit a jump in the circumferential direction.

- H3: The exterior surface of the tunnel shell and the surrounding rock are always in *contact*, i.e., no gaps are occurring.
- H4: There is *no stress transfer* between the *ends* of the top heading or the benches and the adjacent rock during respective construction states, e.g., installation of the top heading. This assumption is made because of lack of respective design provisions within the framework of the NATM.
- H5: The LSCs are assumed to have a 1D linear elastic-ideal plastic load-deformation characteristic. In other words, they have a well-defined ultimate load for a large range of (non elastic) deformations [6]. The experimental load-displacement curve for the LSCs used at the Semmering pilot tunnel is depicted in Figure 1(c).

**Monitoring equipment at the Semmering pilot tunnel:** Five devices for displacement measurements are installed at each measurement cross-section. The positions of the measurement points  $MP1$  to  $MP5$  are defined by the angles  $\varphi_{MP1}$  to  $\varphi_{MP5}$  (see Figure 2(a)). The displacement vectors obtained at these measurement points are referred to as  $\bar{\mathbf{u}}_{MPi}$  with  $i = 1, \dots, 5$ . In addition to the displacements at  $MP1$  to  $MP5$ , the shortenings of the three LSCs were recorded *in situ*. They are referred to as  $\Delta\bar{u}_{\varphi, LSCk}$  with  $k = 1, 2, 3$ . The positions of the LSCs are defined by the angles  $\varphi_{LSC1}$  to  $\varphi_{LSC3}$ , see Figure 2(a).

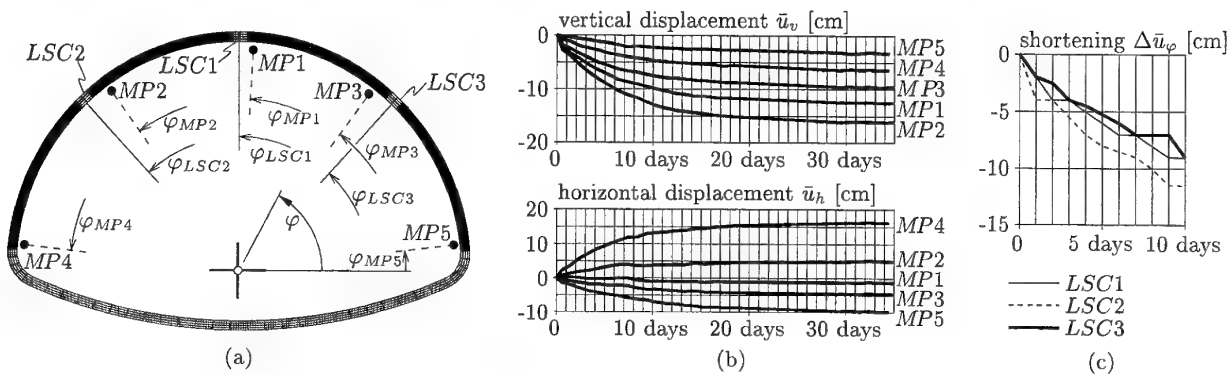


Figure 2: Measured displacements in the Semmering pilot tunnel at km 4.274: (a) location of measurement points (MP) and Lining Stress Controllers (LSC), (b) vertical and horizontal displacements at measurement points, and (c) shortening of LSCs

**Interpolation in time:** The displacements at the measurement points,  $\bar{\mathbf{u}}$ , and the shortening of the LSCs,  $\Delta\bar{u}_{\varphi}$ , are obtained at discrete time values  $t_n$ . For the numerical simulation, the definition of a continuous displacement history, i.e.,  $\bar{\mathbf{u}} = \bar{\mathbf{u}}(t)$  and  $\Delta\bar{u}_{\varphi} = \Delta\bar{u}_{\varphi}(t)$ , is required. This history is determined by means of linear interpolation between the time instants  $t_n$  at which the measurements are performed:

$$\bar{\mathbf{u}}(t) = \bar{\mathbf{u}}(t_n) + \frac{\bar{\mathbf{u}}(t_{n+1}) - \bar{\mathbf{u}}(t_n)}{t_{n+1} - t_n}(t - t_n) \quad \text{and} \quad \Delta\bar{u}_{\varphi}(t) = \Delta\bar{u}_{\varphi}(t_n) + \frac{\Delta\bar{u}_{\varphi}(t_{n+1}) - \Delta\bar{u}_{\varphi}(t_n)}{t_{n+1} - t_n}(t - t_n), \quad (1)$$

for  $t \in [t_n; t_{n+1}]$ .

**Interpolation in space:** According to the structural hypothesis H2, smooth displacement fields are assumed in the parts of the shell made of shotcrete. At the LSCs, the distribution of the radial displacement component  $\bar{u}_r$  is assumed to be smooth as well. In the circumferential direction, however, buckling of the LSCs results in a jump in the distribution of the circumferential displacement  $\bar{u}_{\varphi}$ ,  $\Delta\bar{u}_{\varphi}$ . These jumps are considered by means of Heavyside functions  $H_k$ ,

$$H_k(\varphi) = \begin{cases} 0 & \dots & \text{for } \varphi < \varphi_{LSCk}, \\ 1 & \dots & \text{for } \varphi \geq \varphi_{LSCk}, \end{cases} \quad (2)$$

in the approximation of the circumferential displacement fields at the exterior boundary of the investigated part of the tunnel shell. Altogether, six quadratic functions are employed for this approximation, see Figure 3, namely

$$\bar{u}_{\varphi}^A(\varphi, t) = a_{\varphi}^A(t) + b_{\varphi}^A(t)\varphi + c_{\varphi}^A(t)\varphi^2 + \sum_{k=1}^3 \Delta\bar{u}_{\varphi, LSCk} H_k(\varphi) \quad \text{with } A=I, \dots, VI. \quad (3)$$

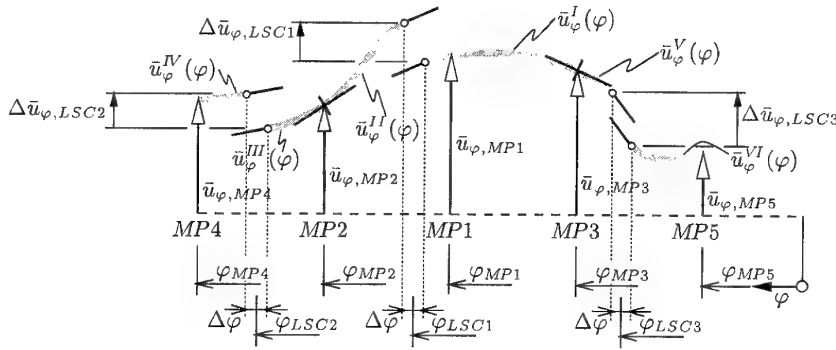


Figure 3: Illustration of the quadratic interpolation of  $\bar{u}_{\varphi}$  on the basis of displacements at measurement points  $MP1$  to  $MP5$ , displacement jumps and prescribed values of  $\bar{u}_{\varphi}' = \partial \bar{u}_{\varphi} / \partial \varphi$  at  $LSC1$  to  $LSC3$

The coefficients  $a_{\varphi}^A$ ,  $b_{\varphi}^A$ , and  $c_{\varphi}^A$  are determined such

- that the displacement values at the measurement points coincide with the respective values measured there, i.e.  $\bar{u}_{\varphi}(\varphi = \varphi_{MPi}) = \bar{u}_{MPi}$ ,
- that continuity of  $\bar{u}_{\varphi}' = \partial \bar{u}_{\varphi} / \partial \varphi$  at  $MP2$  and  $MP3$  is enforced, and
- that the value of  $\bar{u}_{\varphi}'$  at the longitudinal gaps guarantees the agreement of the circumferential axial force  $n_{\varphi}$  in the shotcrete adjacent to the gap with the force in the respective LSC, calculated from the measured shortenings on the basis of structural hypothesis H5.

The last condition requires an additional iterative procedure during the respective non-linear chemomechanical Finite Element analysis.

In contrast to the circumferential direction, the radial displacement component  $\bar{u}_r$  is continuous. Three quadratic functions are used for the interpolation between the values at the measurement points (see Figure 4):

$$\bar{u}_r^A(\varphi, t) = a_r^A(t) + b_r^A(t)\varphi + c_r^A(t)\varphi^2 \quad \text{with } A=I, \dots, III. \quad (4)$$

Continuity of  $\bar{u}_r' = \partial \bar{u}_r / \partial \varphi$  is enforced at  $MP2$  and  $MP3$ .

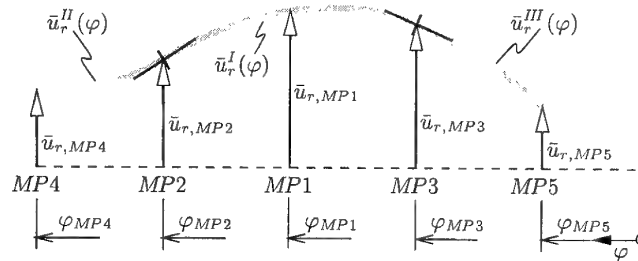


Figure 4: Illustration of the quadratic interpolation of  $u_r$  on the basis of displacements at measurement points  $MP1$  to  $MP5$

In the context of the hybrid method,  $\bar{u}_{\varphi}$  and  $\bar{u}_r$  are prescribed on the outer surface of the tunnel shell. The evaluation of  $\bar{u}_{\varphi}$  and  $\bar{u}_r$  from interpolation of quantities measured in the interior side of the tunnel shell requires the validity of the structural hypothesis H1.

**Material model for shotcrete:** Shotcrete is modelled in the framework of chemically reactive porous media, for a detailed description, see [2] and references therein. Dissipative phenomena at the microlevel of the material are accounted for by means of (internal) state variables and energetically conjugated thermodynamic forces, related to the state variables via state equations. The rates of the internal state variables are related to the corresponding thermodynamic forces by means of evolution equations.

In shotcrete, four such dissipative phenomena govern the material behavior:

(I) The chemical reaction between water and cement, the *hydration* leads to chemical shrinkage strains, ageing elasticity, and strength growth (chemomechanical couplings), furthermore to latent heat production (thermochemical coupling). The state of the reaction is described by the mass per unit volume of the reaction products called hydrates,  $m$ .

(II) *Microcracking* of the hydrates results in permanent or plastic strains  $\varepsilon^p$ . The state of microstructural changes resulting from microcracking (i.e., hardening/softening) is described by hardening variables  $\chi$ , as in classical plasticity theory.

(III) Stress-induced *dislocation-like processes* within the hydrates result in flow (or long-term) creep strains  $\varepsilon^f$ . The state of respective microstructural changes is described by the viscous flow  $\gamma$ .

(IV) Stress-induced *microdiffusion of water* in the capillary pores between the hydrates result in viscous (or short-term) creep strains  $\varepsilon^v$ .

### 3. Chemomechanical Analysis of the Semmering Pilot Tunnel

Figure 5 shows first results of the chemomechanical Finite Element analysis of the tunnel cross section at km 4.274 on the basis of the hybrid method adapted for tunnels in squeezing rock: The LSCs are exclusively loaded in the compressive regime. In the shotcrete part of the tunnel shell, also tensile forces occur. Figure 5(b) reflects the pronounced creep capabilities of shotcrete, whereas the distribution of the circumferential axial force depicted in Figure 5(c) seems to stem from highly non-homogeneous conditions in the adjacent rock. A more detailed interpretation of such results, including the remarkable *bending* of the investigated shell, is a topic of current research activities.

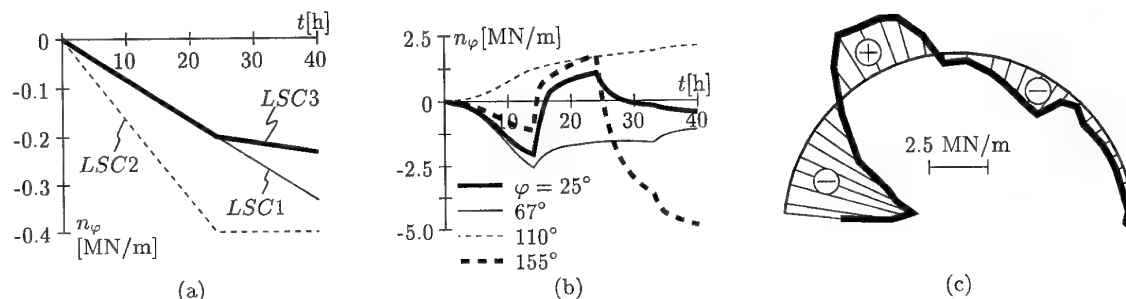


Figure 5: Semmering pilot tunnel: evolution of axial force  $n_\varphi$  in (a) Lining Stress Controllers and (b) four different shell sections; (c) distribution of axial force at  $t = 36$  h

### Acknowledgements

The authors would like to thank W. Schubert and B. Moritz from the Institute for Rock Mechanics and Tunneling of Graz University of Technology for the transfer of measurement data and for fruitful discussions.

### 4. References

- 1 MCCUSKER, T.G.: Soft ground tunneling. In J.O. Bickel, T.R. Kuesel, editors, *Tunnel engineering handbook*, 1982.
- 2 HELLMICH, CH.; LECHNER, M.; LACKNER, R.; MACHT, J.; MANG, H.A.: Creep in shotcrete tunnel shells. In S. Murakami and N. Ohno, editors, 5th IUTAM Symposium on Creep in Structures, Nagoya, Japan, 2000.
- 3 HELLMICH, CH.; ULM, F.J.; MANG, H.A.: Hybrid method for quantification of stress states in shotcrete tunnel shells: combination of 3D *in-situ* displacement measurements and thermochemoplastic material law. In W. Wunderlich, editor, *CD-ROM Proceedings of the European Conference of Computational Mechanics*, Munich, Germany, 1999.
- 4 MACHT, J.; HELLMICH, CH.; MANG, H.A.: Hybrid analyses of shotcrete tunnel shells in squeezing rock. In *CD-ROM Proceedings of the IASS-IACM 2000 Colloquium on Computation of Shells & Spatial Structures*, Chania - Crete, Greece, 2000.
- 5 ROKAHR, R.B.; ZACHOW, R.: Ein neues Verfahren zur täglichen Kontrolle der Auslastung einer Spritzbetonschale. *Felsbau*, 15(6):430-434, 1997. In German.
- 6 SCHUBERT, W.; MORITZ, B.: Controllable ductile support system for tunnels in squeezing rock. *Felsbau*, 16(4):224-227, 1998.

Adresse: DIPL.-ING. JÜRGEN MACHT, DR. TECHN. ROMAN LACKNER, DR. TECHN. CHRISTIAN HELLMICH, PROF. DR. TECHN. DR. H.C. HERBERT MANG, PH.D., Institute for Strength of Materials, Vienna University of Technology, Karlsplatz 13, A-1040 Vienna, Austria

OETTL, G.; STARK, R. F.; HOFSTETTER, G.

## Verification of a Fully Coupled FE Model for Tunneling under Compressed Air

*This paper deals with the verification of a fully coupled finite element model for tunneling under compressed air. The formulation is based on mixture theory treating the soil as a three-phase medium with the constituents: deformable porous soil skeleton, water and air. Starting with a brief outline of the governing equations results of numerical simulations of different laboratory tests and of a large-scale in-situ test are presented and compared with experimental data.*

### 1. Introduction

For tunneling below the groundwater table deformations of the ground and surface settlements are caused both by dewatering of the soil and by the advance of the tunnel face. The application of compressed air for the dewatering process in the vicinity of the tunnel face is motivated by the objective to get smaller ground settlements than by lowering the groundwater table by means of pumping wells and driving the tunnel under atmospheric conditions.

The aim of the current research project is the development and the application of a three-dimensional numerical model for tunneling below the groundwater table taking into account compressed air as a means for dewatering the soil in the vicinity of the tunnel face.

Basically, there exist two different solution strategies for the numerical simulation of tunneling under compressed air. In the first one, an uncoupled numerical approach, the flow of water and compressed air in the soil and the deformations of the soil skeleton are treated in two consecutive steps resulting in a neglect of interactions between the fluid flow and the deformations. The coupled solution procedure applied in the present model permits consideration of the intrinsic coupling of the process of dewatering with the deformations of the soil, thus allowing to properly take into account the interactions of the flow of water and compressed air in the soil with the deformations of the soil skeleton in a physically consistent manner.

### 2. Short outline of the governing equations

A basic constituent of a coupled numerical approach for tunneling under compressed air is a model for the soil, treating the soil as a three-phase medium consisting of the deformable porous soil skeleton and the fluid phases water and compressed air. The mathematical description of the problem is based on mixture theory together with so-called averaging procedures. A soil element can be thought of as a mixture of the three phases soil grains, water and air, which continuously fill the entire volume according to their percentile share.

Following [1] the governing equations for such a three-phase model are described briefly in the subsequent part of this section.

The equilibrium equations are formulated for quasi-static conditions in terms of averaged values for the density  $\bar{\rho}$  and the stresses  $\bar{\sigma}$  of the three-phase mixture as

$$\text{div} \bar{\sigma} + \bar{\rho} \mathbf{g} = \mathbf{0} , \quad (1)$$

where  $\mathbf{g}$  denotes the vector of gravitational acceleration. The total stress tensor  $\bar{\sigma}$  is decomposed into the effective stress tensor  $\bar{\sigma}'$  of the soil skeleton and the hydrostatic stresses  $p^w$  and  $p^a$  of the two fluid phases water and air (using the porosity  $n$  of the soil and the degrees of saturation  $S^w$  and  $S^a$  of the fluids).

The mass balance equation for a fluid phase  $f$  can be derived as

$$\rho^f S^f \dot{\epsilon}_{vol} + n S^f \dot{\rho}^f + n \rho^f \dot{S}^f = -\rho^f \text{div} \tilde{\mathbf{v}}^{fr} , \quad (2)$$

$\tilde{\mathbf{v}}^{fr}$  being Darcy's artificial velocity of the fluid relative to the soil skeleton. Equation (2) reflects the fact that the inflow of fluid mass into a given control volume, represented by the term on the right hand side, can be stored in the volume element either by an increase of the volumetric strain of the soil skeleton, by an increase of the density

of the fluid phase or by an increase of the degree of saturation of the respective fluid phase.

The well-known kinematic relations between the displacements of the soil skeleton and the strains in the soil skeleton are restricted to small displacements and small strains which seems to be a justified assumption for tunneling under compressed air.

To define the material behavior of the various constituents of the model constitutive equations for the soil skeleton, the fluid phases water and air and the capillary stress, which is defined as the difference between the hydrostatic stresses in the air- and the water phase, are required.

### 3. Implementation into a finite element program

For the numerical solution of problems involving a three-phase medium, weak formulations of the equilibrium equations for the three-phase mixture (1) and of the mass balance equations for the fluid phases (2) are required, which can be derived multiplying the respective equations by virtual displacements or virtual hydrostatic stresses and subsequent integration over the domain under consideration. The application of standard finite element procedures together with the implicit and unconditionally stable Euler backward method for the numerical integration in the time domain yields the following coupled set of equations, which has to be solved for the incremental nodal values of the displacements of the soil skeleton and the hydrostatic stresses in the two fluid phases water and air.

$$\begin{bmatrix} \mathbf{K} & \mathbf{C}_{n+1}^w & \mathbf{C}_{n+1}^a \\ (\mathbf{C}_{n+1}^w)^T & -\mathbf{S}_{n+1}^w - \Delta t_{n+1} \mathbf{H}_{n+1}^w & \mathbf{C}_{n+1}^{wa} \\ (\mathbf{C}_{n+1}^a)^T & \mathbf{C}_{n+1}^{wa} & -\mathbf{S}_{n+1}^a - \Delta t_{n+1} \mathbf{H}_{n+1}^a \end{bmatrix} \begin{Bmatrix} \Delta \mathbf{U}_{n+1} \\ \Delta \mathbf{P}_{n+1}^w \\ \Delta \mathbf{P}_{n+1}^a \end{Bmatrix} = \begin{Bmatrix} \Delta \mathbf{f}_{n+1}^{ex} - \mathbf{C}_{n+1}^w \mathbf{P}_n^w - \mathbf{C}_{n+1}^a \mathbf{P}_n^a \\ \Delta t_{n+1} (\dot{\mathbf{f}}_{n+1}^w + \mathbf{H}_{n+1}^w \mathbf{P}_n^w) \\ \Delta t_{n+1} (\dot{\mathbf{f}}_{n+1}^a + \mathbf{H}_{n+1}^a \mathbf{P}_n^a) \end{Bmatrix} \quad (3)$$

The matrices involved in equation (3) are the stiffness matrix  $\mathbf{K}$  of the soil skeleton, the coupling matrices  $\mathbf{C}^w$ ,  $\mathbf{C}^a$  and  $\mathbf{C}^{wa}$  between the different constituents, the compressibility matrices  $\mathbf{S}^w$  and  $\mathbf{S}^a$  and the permeability matrices  $\mathbf{H}^w$  and  $\mathbf{H}^a$ .  $\Delta \mathbf{U}$ ,  $\Delta \mathbf{P}^w$  and  $\Delta \mathbf{P}^a$  denote the global vectors of the incremental nodal values for the displacements of the soil skeleton and for the hydrostatic stresses in the fluid phases. The vector on the right hand side contains the increment of the external forces  $\Delta \mathbf{f}^{ex}$  of the soil skeleton and the flow terms  $\dot{\mathbf{f}}^w$  and  $\dot{\mathbf{f}}^a$ .

Since all the matrices in (3) depend on the nodal values in some form, this system of equations is non-linear (even in the case of linear elastic material behavior of the soil skeleton). Thus, it has to be solved iteratively for each time step, e. g., by means of a direct iteration procedure.

### 4. Verification of the three-phase formulation

The complex implementation procedure of the three-phase soil model into the FE program AFENA [2] was subdivided into several steps. Each of them was checked by the numerical simulation of appropriate experiments.

The first step of the implementation contained the special case of a two-phase formulation taking into account a compressible fluid phase. This subset of the three-phase model was verified by the numerical simulation of a footing on a water saturated soil layer [3] and of a laboratory test (Fig. 1(a)) conducted at the Institute for Soil Mechanics and Foundation Engineering at the Technical University of Graz within the framework of the *Austrian Joint Research Initiative on Numerical Simulation in Tunneling*. In this experiment the loss of compressed air through cracks in the shotcrete lining and the flow of compressed air in the adjacent soil were determined. During the test the time-dependent development of the air pressure in the soil was measured at selected points [4].

Fig. 1(b) shows the computed distribution of the hydrostatic stress in the air phase after steady state conditions have been attained, whereas the dots refer to the respective measured values. The experimental set-up together with numerical results are documented in some more detail in [5].

For the verification of the complete three-phase formulation a laboratory test performed by Liakopoulos [6] was chosen as a first example. The problem deals with dewatering of a sand column of 1 m height under atmospheric conditions, i.e. the dewatering process is driven by gravitation only. Prior to the start of the experiment steady state conditions are prevailing in the sense that water is continuously added at the top and freely drains at the bottom of the column. At the onset of the experiment the water supply is stopped.

Fig. 2(a) shows the transient stresses in the water phase versus height of the sand column for different points of time. Quite substantial differences between the numerical and the experimental data at the beginning of the test (see lines according to 10 minutes time) can be recognized. However, these discrepancies can also be noticed in the

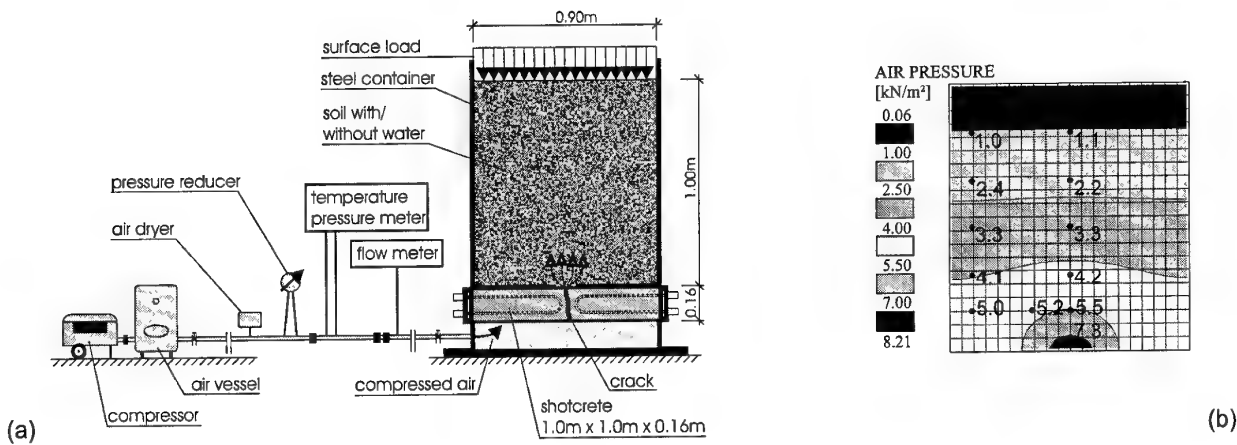


Figure 1: Air flow experiment TU-Graz: (a) Test set-up; (b) Air pressure distribution

computed results for this test, presented by other research groups [1]. Nevertheless, the computed solution agrees fairly well with the measured data after two hours time. Re-running the problem assuming a rigid soil skeleton yields even larger discrepancies in the early test phase indicating that the numerical results could probably be improved using a more refined constitutive law for the behavior of the soil skeleton. Unfortunately, Liakopoulos did not measure any mechanical properties of the soil. As can be seen from Fig. 2(b) there is a sharp decrease of the excess air pressure with a clear peak at about that level of the column where the degree of water saturation of the mixture starts decreasing below fully saturated conditions. Taking a look at the degree of water saturation (Fig. 2(c)) the different lines indicate the position of the de-wetting front moving down the sand column with increasing time. Together with this de-wetting front the peak in the excess hydrostatic air stress is also moving downwards. Since the degree of water saturation decreases, starting from the top of the column, air is allowed to enter the pores of the soil resulting in an increase of the air pressure up to the atmospheric one. Calculations dealing with a time period beyond that of the experiment show that the excess air stresses completely disappear after some time which means that at the final stage atmospheric pressure prevails in the whole soil column again. With respect to the degree of water saturation the numerical simulation up to steady state conditions reveals that the column remains fully saturated at the bottom whereas at the top the degree of water saturation is reduced to about 90 %. This means that for the particular soil considered here it is not possible to completely dewater the specimen under the influence of gravitation only.

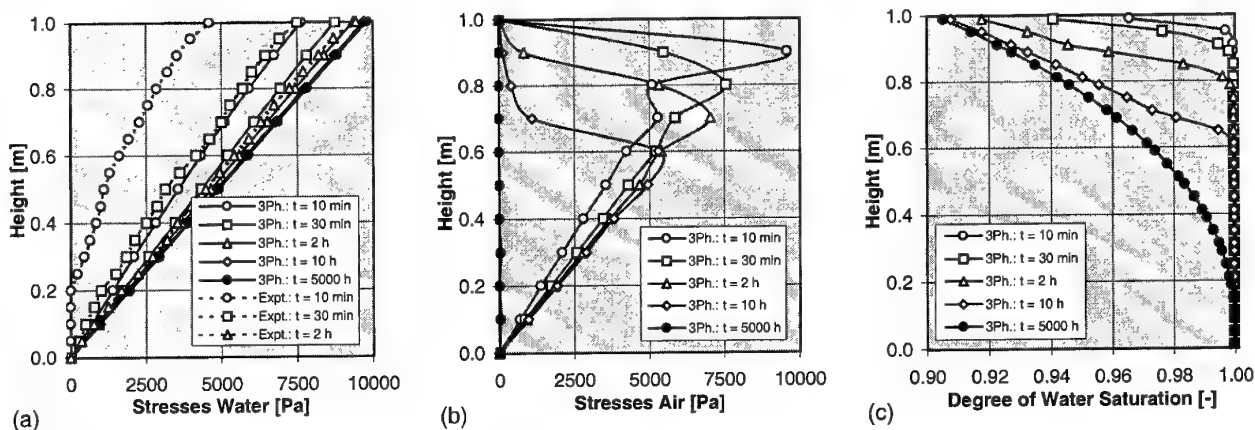


Figure 2: Liakopoulos experiment: Stresses in (a) the water phase and (b) the air phase; (c) Water saturation

In connection with the application of compressed air at the subway construction site in Essen, Germany, a full scale in-situ air permeability test was carried out [7]. It is considered as a second example for the verification of the complete three-phase formulation. Fig. 3(a) shows a sketch of the experimental set-up. In order to investigate the air permeability of the Essen soil and to study the influence of the flow of compressed air on the deformations of the soil skeleton, three sets of tests were performed. In these tests compressed air was injected at different depths

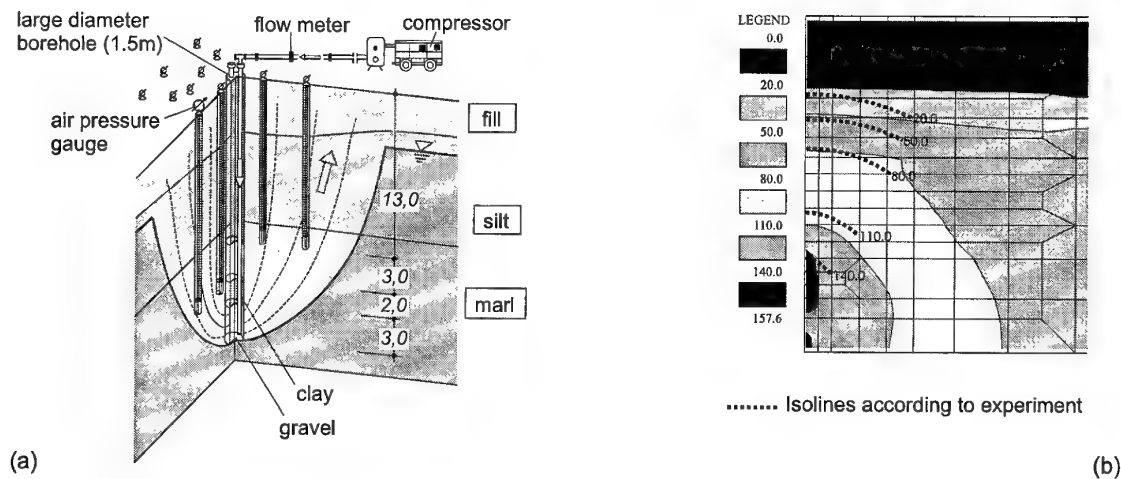


Figure 3: Air permeability test Essen: (a) Test set-up; (b) Air pressure distribution

below the ground surface. In the experiment considered here compressed air was pumped into the ground via a bore hole between a depth of 18 and 21 m below the ground surface with an excess air pressure up to 2.35 bar. This air pressure was applied in three steps keeping the pressure constant for about one day at each of the three level.

Fig. 3(b) finally shows the computed air pressure distribution at the end of the first pressure level of 1.60 bar together with the isolines composed from the measured values of the air pressure in the experiment. The agreement between the measured and the computed air pressure is remarkably good.

## 5. Conclusions

Apart from a given short summary of the theoretical framework of a three-phase soil model for tunneling under compressed air and of the different steps of the implementation into a FE code, the paper focussed on the verification of the developed model. Comparisons of the numerical results of the two-phase- as well as of the complete three-phase formulation with both experimental data and solutions reported in the literature show fairly good agreement.

## Acknowledgements

The research project, described in this paper, is funded by the Austrian Science Fund (FWF) under project number S08005-TEC. This support is gratefully acknowledged.

## 6. References

- 1 LEWIS, R.W., SCHREFLER, B.A.: The Finite Element Method in the Static and Dynamic Deformation and Consolidation of Porous Media. John Wiley & Sons, Chichester, 1998.
- 2 CARTER, J.P., BALAAM, N.P.: AFENA User's Manual, Version 5.0. Centre for Geotechnical Research, University of Sydney, Sydney, Australia, 1995.
- 3 HOFSTETTER, G., OETTL, G., STARK, R.F.: Development of a three-phase soil model for the simulation of tunnelling under compressed air. *Felsbau*. **17**, No. 1 (1999), 26–31.
- 4 KAMMERER, G., SEMPRICH, S.: The prediction of the air loss in tunnelling under compressed air. *Felsbau*. **17**, No. 1 (1999), 32–35.
- 5 HOFSTETTER, G., OETTL, G., STARK, R.F.: A soil model for tunneling under compressed air. In: WUNDERLICH, W. (ed): Proceedings of the European Conference on Computational Mechanics, ECCM '99. Lehrstuhl für Statik, TU München, München, 1999, CD-Rom.
- 6 LIAKOPOULOS, A.C.: Transient Flow Through Unsaturated Porous Media. PhD thesis, University of California, Berkeley, 1965.
- 7 KRAMER, J., SEMPRICH, S.: Erfahrungen über Druckluftverbrauch bei der Spritzbetonbauweise. *Taschenbuch für den Tunnelbau*. **13** (1989), 91–153.

Addresses: DIPL.-ING. GERHARD OETTL, ASS. PROF. DR. RUDOLF STARK, UNIV. PROF. DR. GUENTER HOFSTETTER, University of Innsbruck, Institute for Structural Analysis and Strength of Materials, Technikerstraße 13, A-6020 Innsbruck, Austria.



SCHWEIGER, H.F., SCHULLER, H.

## Multilaminate Models for Numerical Simulation of Tunnel Excavations According to NATM

*Excavation of tunnels with low overburden under unfavourable geological conditions can lead to extensive shear band formation in the ground which can even cause the collapse of the structure. In this paper a constitutive model embedded within the Multilaminate framework is presented which is able to capture the mechanism of strain localisation and the formation of shear bands. The model includes deviatoric hardening in the pre-peak range and strain softening in the post-peak behaviour. Application to numerical simulations of a biaxial test and a shallow NATM tunnel excavation in staged construction sequence demonstrates the potential of the model to predict shear band formation and possible failure mechanisms with sufficient accuracy for practical problems.*

### 1. Introduction

For excavation of tunnels with low overburden, a number of geotechnical problems have to be solved. An important role play strain localisation phenomena such as the formation of shear bands in the ground. During strain localisation, the nearly homogeneous deformation behaviour of a body is abruptly changed into highly concentrated deformation patterns and a single or multiple shear bands or cracks form.

Extensive shear band formation under unfavourable geological conditions as well as an inappropriate construction sequence may lead to significant deformations of the tunnel structure and could even cause a collapse. A numerical model, which is able to simulate the deformation behaviour of a structure near failure and predict possible failure mechanisms could, therefore, contribute to a safer and more economic design of tunnels with low overburden.

In this contribution, a constitutive model embedded within the Multilaminate framework is applied to finite element analyses of a shallow tunnel excavation. Deviatoric hardening plasticity in the pre-peak range and a simple strain softening formulation to describe the post-peak behaviour are incorporated. The tunnel excavation according to the principles of NATM (New Austrian Tunnelling Method) was simulated by modelling a staged construction sequence with excavation of top heading, bench and invert.

### 2. Multilaminate Concept and Constitutive Relations

The Multilaminate Model was developed for application in rock engineering by Zienkiewicz & Pande [12]. Subsequently, it has been extended for modelling clays [5] and sands [7], [4]. Recent developments include research on strain localisation in dense sands [3]. A similar approach to describe the behaviour of concrete and soils is the microplane model developed by Bažant et al. [1].

The basic idea of the Multilaminate concept is that a block of soil material is thought to be intersected by an infinite number of randomly distributed planes. It is assumed that the deformation behaviour of this soil can be obtained from sliding along some of the planes under effective normal and shear stresses. These assumptions are transferred into the finite element model where a certain number of integration planes is located within each integration point (Gauss point). Number and orientation of these integration planes either follow integration rules (3-D case) or they are evenly distributed about a common axes (2-D plane strain case).

The stress state in the integration point is transformed into normal and shear stresses on each integration plane and, consequently, evaluated by a yield function and plastic potential function. The resulting plastic strains of all planes are integrated to obtain the behaviour in the respective integration point.

The constitutive relations are formulated on each integration plane individually, thus, directional anisotropy can be achieved in a natural and physically meaningful manner depending on current stress state and load history. As a yield criterion, Mohr-Coulomb's failure criterion was enhanced by introduction of a mobilised friction angle  $\varphi_m$ . Mobilisation of the friction angle is governed by accumulated plastic shear strains on each integration plane

individually (deviatoric hardening) until the peak friction angle is reached. With increasing load, shear strains concentrate on a few planes as demonstrated in [8] and [10]. When the peak friction angle is reached on some of the integration planes and shearing continues, the model switches into strain softening and friction angle and cohesion are reduced on the respective planes to residual values.

A non-associated flow rule employing a mobilised dilation angle  $\psi_m$  governs plastic flow and dilatancy. The mobilised dilation angle depends on the mobilised friction angle  $\phi_m$  according to the stress-dilatancy theory proposed by Rowe [6].

The Multilaminate Model was implemented into the finite element code BEFE [2] which solves nonlinear equation systems utilising a viscoplastic algorithm. A detailed description of the implemented yield function and plastic potential function are included in [9].

### 3. Numerical Simulation of Biaxial Test

A biaxial test (plane strain compression) was modelled to demonstrate the potential of the proposed method in capturing shear band formation. In addition, four different meshes consisting of 126, 224, 504 and 880 elements, respectively, were analysed to evaluate possible mesh dependences in the post-peak range. In all calculations, a plane strain formulation of the Multilaminate Model employing 64 integration planes in each integration point was used.

The sample of the dimensions 1:3.4 is vertically constrained at the bottom. The top of the sample is bounded by a horizontally constrained stiff top platen. Full friction between top platen and the sample introduces a slightly inhomogeneous stress distribution which, at a certain load stage, initiates strain localisation. Starting from an initial stress of  $\sigma_1 = \sigma_3 = -100 \text{ kN/m}^2$ , strain controlled loading was modelled by stepwise application of nodal displacements to the top platen.

In Fig.1, curves of the mean vertical stresses  $\sigma_1$  versus vertical strains  $\epsilon_1$  are presented. The curves for all meshes, consisting of 126, 224, 504 and 880 elements, respectively, are congruent up to the peak which is reached at a vertical strain  $\epsilon_1$  of approximately 3.0%. In the post peak softening range some deviation can be observed, namely the meshes consisting of a higher number of elements show slightly more pronounced softening. However, when the residual strength is reached at an approximate vertical strain of  $\epsilon_1 = 4.0\%$  all curves are in a close range.

Fig.2 presents contour lines of the mobilised friction angle  $\phi_m$  for all meshes at residual state ( $\epsilon_1 = 4.0\%$ ). Obviously, in all calculations the mobilised friction angle has been reduced to its residual value in a distinct zone indicated by the dark strip across the sample. This zone represents the shear band. There are some differences in the location and the width of the shear band which depends on the element size.

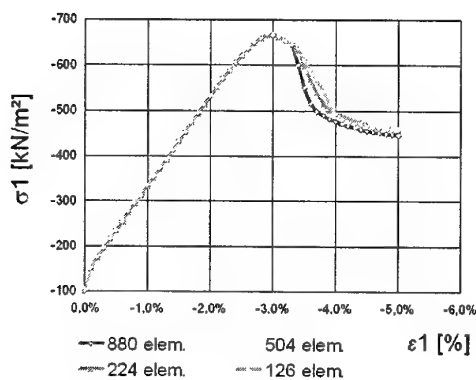


Fig.1: Stress-strain curves for biaxial tests using different meshes

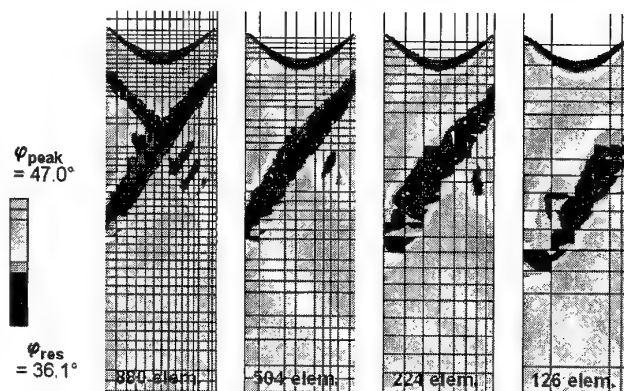


Fig.2: Contour lines of mobilised friction angle  $\phi_m$  for meshes consisting of 880, 504, 224 and 126 elements

#### 4. NATM Tunnel Excavation

The deformation behaviour of tunnel excavations with low overburden usually can not be matched very well if linear elastic – perfectly plastic material models are employed. Utilising more advanced constitutive models prediction capabilities can be significantly improved. However, from a practical point of view the model should be as simple as possible to avoid excessive parameter studies and computational costs [11].

The proposed model was applied to numerical analyses of a practical boundary value problem, namely a tunnel excavation with low overburden. A tunnel excavation according to the principles of the “New Austrian Tunnelling Method” (NATM) was modelled considering staged construction sequence: First, the top heading (the upper part of the tunnel cross section) was excavated and a shotcrete layer was applied to the tunnel walls. Then, construction sequence consisted of excavation of the bench (middle part of the cross section), application of shotcrete to the bench walls, excavation of the invert (lowest part) and, finally, application of the shotcrete layer to the tunnel invert.

The aim of the numerical simulations presented here was to show that a possible failure mechanism due to formation and progressive development of shear bands could be captured with the proposed model. Therefore, the previously described construction sequence was altered at the stage of bench excavation to enable development of extensive shear zones. The shotcrete layer usually stabilising the tunnel was not applied, at the same time the forces due to bench excavation were increased in increments of 5% until failure occurred. During this process, development of shear strains and mobilisation of the friction angle  $\phi_m$  was monitored. After application of 80% of the excavation load no converged solution could be obtained.

In Fig.3 the failure mechanism at bench excavation is demonstrated by showing contour lines of the mobilised friction angle. The shotcrete lining consisting of one line of elements can be recognised at the top heading of the tunnel while it is not installed at the tunnel bench. Thus, with increasing load a failure mechanism forms initiating from the side of the tunnel next to the bench excavation. Light areas in the contour plot stand for zones where the shear resistance of the soil material is fully mobilised ( $\phi_m = \phi_{peak}$ ) or the material behaviour has switched into softening on some integration planes. Consequently, at these planes the friction angle is decreased with increasing shear strains until the residual value of the friction angle is reached. Beginning from the softening zone at the side of the tunnel, a shear band starts growing towards the ground surface. When 80% of the nodal forces due to excavation are applied, the shear band reaches the ground surface. No convergence can be reached and a failure mechanism forms leading to large deformations within the shear band and collapse of the tunnel. This mechanism corresponds well with experimental findings and failure mechanisms which have occurred at real tunnel construction sites.

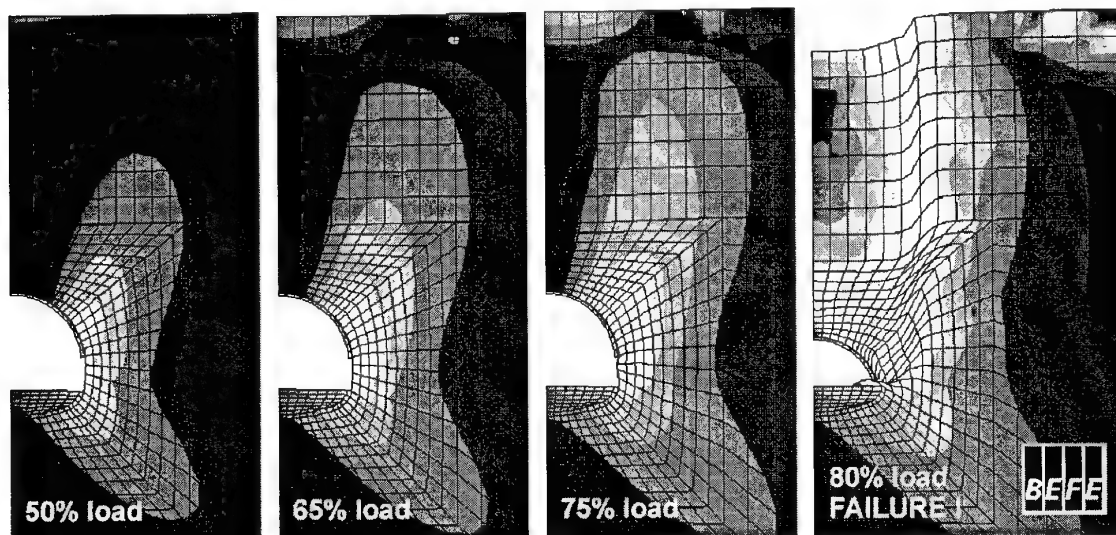


Fig.3: Failure mechanism due to formation of shear band for shallow tunnel excavation, contour lines of mobilised friction angle

## 5. Conclusions

It was demonstrated that the deformation behaviour involving strain localisation and the progressive formation of shear bands can be captured in a realistic way with the proposed model. The constitutive model utilises a strain hardening / strain softening formulation within the Multilaminate framework. Numerical simulations of biaxial tests show that mesh dependences in the post-peak softening range can be mastered with sufficient accuracy from a practical point of view. Well predictions of a possible collapse mechanism are obtained for the numerical analysis of a shallow tunnel excavation with staged construction sequence (top heading / bench / invert).

## Acknowledgements

*This research project is supported by the Austrian Science Fund FWF (project P12577-TEC) and associated with the Joint Research Initiative "Numerical Simulation in Tunnelling" SITU.*

## 6. References

- 1 BAZANT, Z.P., PRAT, P.C.: Microplane model for brittle-plastic material: I. theory; Journal of Engineering Mechanics, Vol.114 / No.10 (1988), 1672-1688.
- 2 BEER, G.; BEFE user's and reference manual; CSS, Carnerig. 10, Graz, Austria (1999).
- 3 KARSTUNEN, M.: Numerical modelling of strain localization in dense sands; Acta Polytechnica Scandinavica No. 113, The Finnish Academy of Technology, Espoo (1999).
- 4 KRAJEWSKI, W.: Mathematisch-numerische und experimentelle Untersuchungen zur Bestimmung der Tragfähigkeit von in Sand gegründeten, vertikal belasteten Pfählen; Veröffentlichungen des Instituts für Grundbau, Bodenmechanik, Felsmechanik und Verkehrswasserbau der RWTH Aachen, Heft 13 (1986).
- 5 PANDE, G.N., SHARMA, K.G.: Multi-laminate model of clays – a numerical evaluation of the influence of rotation of the principal stress axes; Int. J. Num. Anal. Meth. Geomech. 7 (1983), 397-418.
- 6 ROWE, P.W.: Theoretical meaning and observed values of deformation parameters for soil. Stress strain behaviour of soils; Proc. Roscoe Memorial symp., Foulis, Henley-on-Thames (1972), 143-194.
- 7 SADRNEJAD, S.A., PANDE, G.N.: A multilaminate model for sands; Proc. 3<sup>rd</sup> Int. Symp. Num. Models Geomech., Elsevier Applied Science, London (1989), 17-27.
- 8 SCHULLER, H., SCHWEIGER, H.F.: Application of multilaminate model for shallow tunnelling; Felsbau 17, Nr.1 (1999), 44-47.
- 9 SCHULLER, H., SCHWEIGER H.F.: Numerical modelling of tunnel excavation in soils using the multilaminate model; Proc. Numerical Models in Geomechanics NUMOG VII Graz, Balkema, Rotterdam (1999), 341-346.
- 10 SCHWEIGER, H.F., KARSTUNEN, M., PANDE, G.N.: Modelling strain localization in anisotropically consolidated samples; Proc. 6<sup>th</sup> Int. Symp. Num. Models Geomech., Balkema, Rotterdam (1997), 181-186.
- 11 SCHWEIGER, H.F., KOFLER, M., SCHULLER, H.: Some recent developments in the finite element analysis of shallow tunnels; Felsbau 17, Nr. 5 (1999), 426-431.
- 12 ZIENKIEWICZ, O.C., PANDE G.N.: Time dependent multi-laminate model of rocks – a numerical study of deformation and failure of rock masses. Int. J. Num. Anal. Meth. Geomech. 1 (1977), 219-247.

*Address:* AO.UNIV.PROF. DIPL.-ING. DR. TECHN. HELMUT F. SCHWEIGER, DIPL.-ING. HARTMUT SCHULLER, Institute for Soil Mechanics and Foundation Engineering, Graz University of Technology, Rechbauerstr. 12, A-8010 Graz, Austria.

Website: [http://www.tu-graz.ac.at/geotechnical\\_group](http://www.tu-graz.ac.at/geotechnical_group)

E-mail: [schweiger@ibg.tu-graz.ac.at](mailto:schweiger@ibg.tu-graz.ac.at) ; [schuller@ibg.tu-graz.ac.at](mailto:schuller@ibg.tu-graz.ac.at)

STERPI, D.

## Strain Localisation Effects in the Stability Analysis of Shallow Openings

*The localisation of strains within shear bands represents an important aspect of the behaviour of rocks and dense soils which may govern the stability of relevant structures. The formation of shear bands is often associated to a loss of the material mechanical resistance and stiffness, eventually leading to progressive failure. A numerical approach accounting for strain localisation and for the consequent gradual loss of shear strength is adopted for the simulation of laboratory tests on 2- and 3-dimensional models of shallow tunnels. These tests allow for the validation of the numerical procedure on the basis of a comparison between experimental and numerical results.*

### 1. Introduction

The stability analysis of shallow underground openings presents non negligible problems when dealing with media characterised by strain softening behaviour, such as rock masses, dense sands or overconsolidated clays.

In geotechnical engineering the term "strain softening" customarily denotes materials that, when subjected for instance to a compression test, show a loss of their overall load carrying capacity with increasing deformation, after a peak load level has been reached. Quite often this behaviour is associated with the formation of zones of limited thickness where irreversible strains localise, referred to as shear bands. When dealing with large scale engineering problems, the shear bands represent potential sliding surfaces that could affect the overall stability by originating a collapse mechanism.

When a tunnel is driven into a formation of stiff soil or weak rock, characterised by a low value of the coefficient of earth pressure at rest, zones of localisation initiate at the tunnel springlines, if the opening is not properly supported. Their extension might rapidly increase with the progress of excavation, developing upwards into arched shear bands that eventually join at a point above the tunnel crown, or reach the ground surface in the case of shallow openings. In these cases a collapse mechanism may form with the consequent failure of the tunnel crown. A more complex geometry of the collapse mechanism can be observed in the cross section of parallel tunnels. In this case, in fact, the shear bands might develop on the sides of the tunnels as well as within the ground wall separating them. In addition, if the stability of the tunnel face is not ensured during excavation, the spread of shear bands ahead of the excavation face can lead to the inward sliding movement of an unstable mass of soil.

The attention is focused here on the numerical modelling of strain localisation and softening effects as observed during laboratory tests on small scale models of shallow tunnels. These tests were carried out on both 2D (plane strain) and 3D models, to investigate the stability, respectively in the cross section and at the excavation face, of tunnels driven in frictional media. The calibration of the numerical model accounting for softening effects is discussed on the basis of the experimental results. These allow for the validation of the hypotheses introduced in the proposed model and for the refinement of the numerical procedure.

### 2. Numerical Approach to Strain Localisation and Softening

For geological media, the process of initiation and subsequent spread of the localisation zones seems to depend on two aspects. From a "structural" point of view, the loss of load carrying capacity is originated from a sort of structural instability of the material, which is not specifically related to a loss of shear resistance introduced in the stress-strain relation, but occurs when, for a given increment of the external actions, a bifurcation point is reached, i.e. the uniqueness of the solution of the equations governing the stress analysis problem is lost [3]. On the other hand, from a "constitutive" view point, the strain softening phenomenon can be seen as a consequence of an intrinsic property of the material, shown at a constitutive level as a loss of shear strength with increasing deformation [1]. The associated shear band formation is caused by a progressive local damage of the material, in terms of its mechanical deformability and resistance, depending on the cumulated permanent strains or on the plastic strain energy.

The procedure for the strain softening analysis here adopted accounts for both aspects in considering separately two "phases" of the phenomenon, namely: the onset of localisation, based on a structural approach, and the spread and coalescence of the shear bands, based on a constitutive approach. The two phases correspond to two stages of a non-linear finite

element procedure, implemented in the code SoSIA, for Soil-Structure Interaction Analysis [2].

The first stage consists in checking a particular local condition, at each integration point and at the end of each loading step, the fulfilment of which indicates the local onset of strain localisation. This condition stems from the shear band analysis of bifurcation theory, which detects the occurrence of an alternative solution in the form of a planar discontinuity in the strain field. For a plane strain problem in a 2-dimensional system  $\underline{x}$ , this discontinuity can be kinematically described by the difference  $\Delta \underline{\dot{u}}$  between the displacement rates at its two sides, which is assumed to vary linearly with the distance from the discontinuity itself:

$$\Delta \underline{\dot{u}} = \dot{g} \underline{m} \cdot \underline{n}^T \cdot \underline{x} \quad (1)$$

In the above equation,  $\underline{n}$  is the unit vector normal to the discontinuity,  $\underline{m}$  is the unit vector defining the direction of the difference  $\Delta \underline{\dot{u}}$  and  $\dot{g}$  is its amplitude.

The strain compatibility, in terms of the difference  $\Delta \underline{\dot{\epsilon}}$  between the strain rates at the two sides, and the equilibrium conditions for the stress rate  $\underline{\dot{\sigma}}$  across the discontinuity, can be expressed in the following matrix forms:

$$\Delta \underline{\dot{\epsilon}} = \dot{g} \underline{N} \cdot \underline{m} \quad , \quad \underline{N}^T \cdot \Delta \underline{\dot{\sigma}} = 0 \quad , \quad (2a,b)$$

where matrix  $\underline{N}$  collects the direction cosines of unit vector  $\underline{n}$ .

Considering now an elastic perfectly plastic material and assuming that the material at the two sides exhibits the same mechanical behaviour after the onset of the discontinuity, the following constitutive relationship holds:

$$\Delta \underline{\dot{\sigma}} = \underline{D}^{ep}(\underline{\sigma}) \cdot \Delta \underline{\dot{\epsilon}} \quad , \quad (3)$$

in which  $\underline{D}^{ep}$  represents the tangent elastic plastic constitutive matrix. By substituting eq.(3) into eq.(2b) and taking into account eq.(2a), a homogeneous system of governing equations is arrived at, which admits non trivial solutions  $\dot{g} \neq 0$  if and only if the following so called "condition of localisation" is fulfilled:

$$\det[\underline{N}^T \cdot \underline{D}^{ep}(\underline{\sigma}) \cdot \underline{N}] = 0 \quad . \quad (4)$$

It has been proved that this condition can be met even for perfectly plastic, or positive hardening, materials, if the plastic flow rule is non-associated [5]. Further details and comments on the derivation of this equation can be found in [6].

The second stage of the analysis accounts for the actual coalescence of localisation zones into shear bands and, therefore, it is activated only for the integration points where condition (4) has been fulfilled. Consider that the occurrence of a strain discontinuity causes a change in the local structure of the material: for instance an increase of void ratio or a decrease of relative density due to the dilatancy effects. Consequently, a local loss of shear resistance and stiffness might occur.

This suggests to base the second stage of the analysis on a procedure in which the shear strength and stiffness parameters are gradually reduced, with increasing permanent strains, from their peak to their fully softened values. In the calculations, this reduction is linearly related to the increment of the square root of the second invariant of deviatoric plastic strains, with respect to the corresponding value at the onset of localisation. The rate of this reduction depends on the material mechanical characteristics and on a "measure" of the problem discretisation. It has been shown, in fact, that the solution of a numerical strain softening analysis is dependent on the adopted discretisation, which affects thickness and direction of the computed shear bands. The mesh-dependence is induced by the very nature of the finite element method, since the loss of shear strength and stiffness is evenly distributed over a zone that depends on the size of the elements. As a consequence, the decrease of the material parameters adopted in the calculations should be related also to the finite element size [4]. This provision has been adopted here, by keeping constant the product between the average element size and the rate of reduction of the mechanical parameters.

### 3. Applications

The described approach has been applied to the numerical simulation of laboratory tests on 2D and 3D small-scale models of shallow tunnels.

The first series of 2D, plane strain tests was performed at the Rock Mechanics Laboratory of Kobe University (Japan), using an assemblage of aluminium bars as "analogical" soil [8]. The mechanical characterisation of this material was based on tests performed, with non conventional devices, on bar assemblages having the same relative density adopted during the tunnel tests. The tunnel tests are set up by lying the bars within a rigid frame, where one or two steel cylinders are located, that represent the cross section of one tunnel or two parallel tunnels. The cylinders have a diameter of 15cm and contain pressurised airbags. The excavation process is simulated by removing the steel cylinders and by decreasing the air pressure in subsequent steps. The induced stresses and strains are obtained by embedded load cells and by digitised pictures of the aluminium bar assembly.



In the tests featuring a single tunnel, the collapse was originated by the development of two shear bands from the tunnel springlines towards the surface and a negligible deformation was observed in the mass limited by them (fig.1.a). Shape and thickness of the bands depend on the tunnel depth and on the relative density of the medium. The numerical simulation of these tests showed a progressive spread of a shear band whose position was rather similar to the one experimentally observed (fig.1.b,c).

In the case of two parallel tunnels, the simultaneous reduction of pressure gives rise to a gradual increment of vertical stresses at the springline level, which leads to the failure of the separating wall. Consequently, the load is transferred to the external sides of the tunnels, which then behave as a single wide opening. This additional stress increment leads to the initiation of lateral shear bands and, eventually, to the collapse (fig.2.a). In fig.2.b the numerically evaluated contour lines of the deviatoric plastic strains show a high level of deformation within the soil wall and the presence of a lateral shear band.

It can be also observed that the numerical analyses correctly estimated the internal pressures corresponding to the collapse for both cases of single (fig.3.a) and double tunnel (fig.3.b).

A second series of tests was performed at the Laboratories of Mitsubishi Heavy Industries in Takasago (Japan), concerning a 3D steel model of a horse shoe shaped tunnel [7]. The tunnel model has a crown diameter of 1.2 m and the depth of the crown is 1.3 m. The model is contained within a tank filled with a medium uniform sand, from an alluvial deposit of central Japan. Following a procedure similar to the one used in the previous tests, the stability of the excavation face was investigated by measuring the displacements induced in the sand mass by the gradual reduction of the pressure of an airbag contained within the tunnel model.

Some 2D, plane strain calculations were carried out with reference to the longitudinal section of the tunnel. Although they do not allow for a quantitative comparison between experimental and numerical results, they permit a qualitative estimation of the influence of softening on the overall behaviour of the tunnel. Fig.4.a shows that two shear bands develop from the tunnel crown and from its invert arch and gradually reach the ground surface, in agreement with the experimental observation. On the contrary, if the strain softening effects are neglected in the calculations, the shear strains concentrate only within a limited zone ahead of the tunnel face and the shape of the collapse mechanism is not properly predicted (fig.4.b).

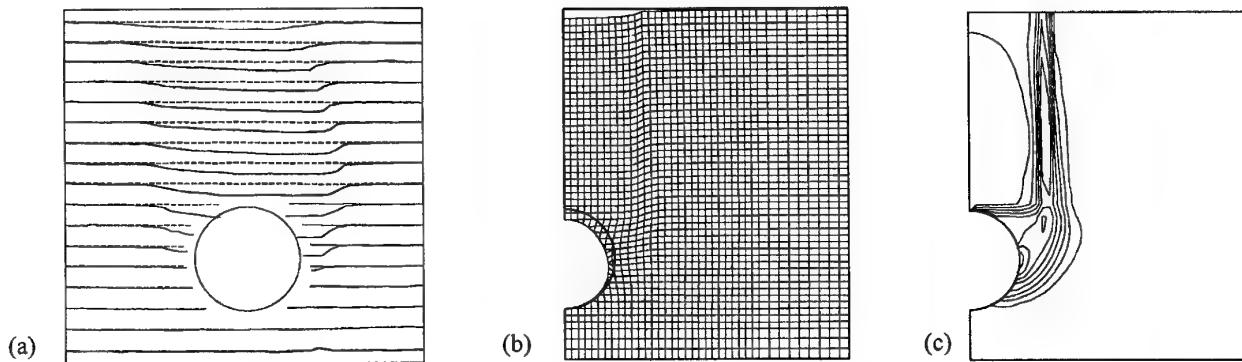


Fig.1. Single tunnel problem: comparison between the experimentally measured settlements (a) and the calculated displacements (b) and deviatoric plastic strains (c) (contour lines of the square root of the 2nd invariant of dev.pl.strains, min=.3%, increment=.6%)

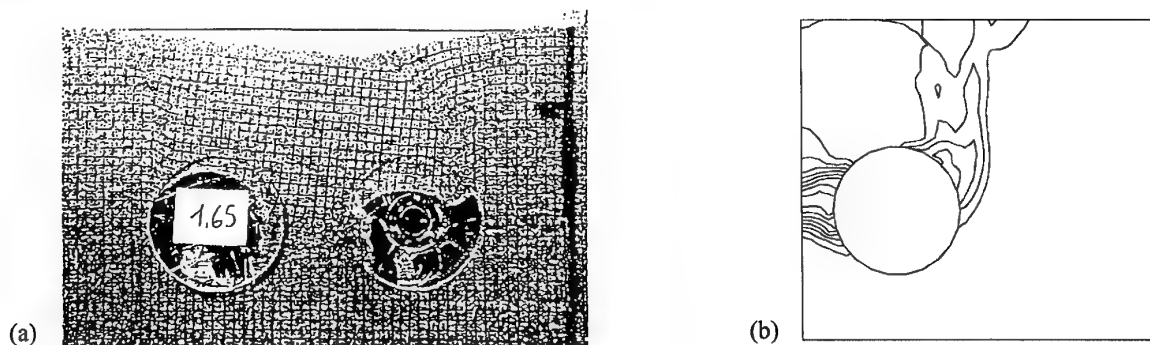


Fig.2. Double tunnel problem: comparison between the experimentally observed collapse (a) and the calculated deviatoric plastic strains (b) (contour lines of the square root of the 2nd invariant of dev.pl.strains, min=.5%, increment=.5%)

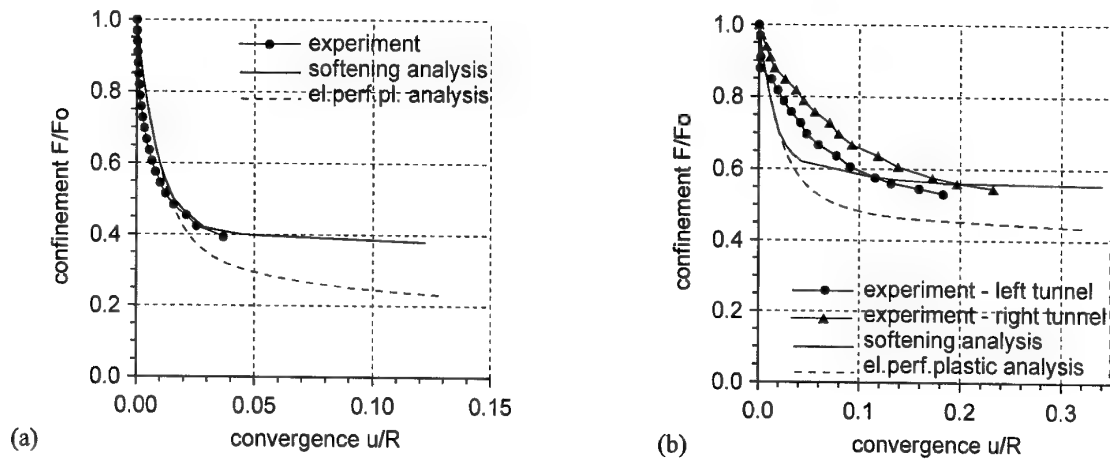


Fig.3. Confinement pressure vs. crown convergence curves for the single (a) and double (b) tunnel problems.

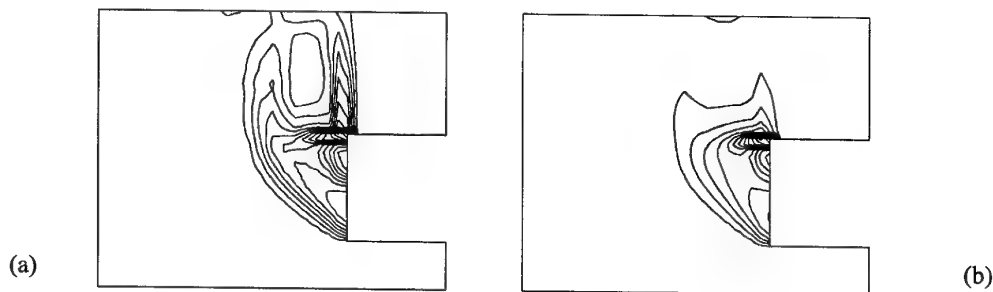


Fig.4. Numerically evaluated deviatoric plastic strains at the excavation face at collapse for strain softening (a) and elastic perfectly plastic (b) analyses (contour lines of the square root of the 2nd invar. of dev.pl.strains, min=.5%, increment=.5%)

#### 4. Conclusions

A finite element model accounting for strain localization and softening effects has been adopted for the numerical interpretation of laboratory tests on 2D and 3D models of tunnels. The procedure, which can be seen as a relatively simple extension of a standard non linear finite element analysis, turned out to be effective in the prediction of the tunnel test results. In particular, collapse load and mechanism suggested by the numerical analyses appear in fairly good agreement with the experimental measurements.

#### Acknowledgements

The financial supports of the Italian Ministry of the University and Research and of the Construction Engineering Research Foundation (Kobe, Japan) are gratefully acknowledged.

#### 5. References

1. BAZANT, Z., BELYTSCHKO, T., CHANG, T.P.: Continuum theory for strain-softening; J. of Eng. Mech. ASCE **110** (1984), 1666-1692.
2. CIVIDINI, A., GIODA, G.: Finite element analysis of direct shear tests on stiff clays; Int.J.Numer.Anal.Meth.Geomech. **16** (1992) 869-886.
3. HILL, R.: A general theory of uniqueness and stability in elastic-plastic solids; J. Mech. Phys. Solids **6** (1958), 236-249.
4. PIETRUSZCZAK, S., MROZ, Z.: Finite element analysis of deformation of strain-softening materials; Int.J.Num.Meth.Eng. **17** (1981) 327-334.
5. RICE, J.R.: The localization of plastic deformation; in KOITER, ed., 14th Int.Conf. Theoretical and Applied Mech.(1976) North Holland.
6. STERPI, D.: An analysis of geotechnical problems involving strain softening effects; Int.J.Numer.Anal.Methods Geomech. **23** (1999), 1427-1454.
7. STERPI, D., CIVIDINI, A., SAKURAI, S., NISHITAKE, S.: Laboratory model tests and numerical analysis of shallow tunnels; in BARLA, ed., Proc. Int. Symp. Prediction and Performance in Rock Mech. and Rock Eng., Eurock'96, (1996) Balkema, 689-696.
8. STERPI, D., SAKURAI, S.: Numerical analysis of laboratory tests on a model tunnel; in ASAKA ET AL., eds., Proc. Int. Symp. Deformation and Progressive Failure in Geomechanics, (1997) Elsevier Science, 757-762.

Address: DR. DONATELLA STERPI, Department of Structural Engineering, Politecnico di Milano, I-20133 Milano, Italy



BRUNS, J.-U.; POPP, K.

## Nonlinearity Tests

*A possible first step in a signal based system identification procedure is a nonlinearity test. The result can be used to decide whether or not simple linear models are sufficient to describe the observed dynamics. It determines if modal analysis or other methods that assume linear system behaviour yield valid results. After a brief overview two tests are discussed in detail and results of an application to test systems are given.*

### 1. Introduction

In many engineering disciplines the identification of dynamical systems from measured signals is an important task in the modeling process. If information on the structure of the system is available, the task is reduced to the identification of parameters. Often, however, either such information is not available or the system structure is known but is very complex and the relevant part of the structure is not easily discerned.

A possible first step in the identification process is to ask whether a simple linear model is sufficient to describe the dynamics or whether more complex nonlinear models have to be considered. The answer to this question does have a strong impact on the experimental procedures used in the identification procedure as well as on possible control schemes to be implemented. Ideally, a nonlinearity test should not only decide whether a system can be regarded as linear or not, but also give an indication about the strength and the type of nonlinearity contained in the system.

### 2. Overview

Table 1 gives an overview of a selection of signal based nonlinearity tests that can be found in the literature. The tests have been grouped according to the domain in which they operate—the time or the frequency domain—and the way the system input is dealt with—some tests make assumptions about the type of input while other tests require that the system input is measured. The overview is by no means complete and the reader is referred to [1, 2, 3, 4] for further references.

The tests by Keenan [5] and Tsay [6] assume the input signal to be a strictly stationary zero mean random process, the null hypothesis of a linear system is accepted or rejected based on a statistic calculated from multiple regression steps. In Surrogate Data Tests [7] measured signals are modified (surrogate data) in a way that is consistent with the null hypothesis of a linear system. The second step is to compute a statistic that is capable of discriminating linear and nonlinear systems. The Method of Internal Harmonics Cross-Correlation by Dimentberg, Sokolov and Haenisch [8, 9] is applicable to systems that are excited by broadband random signals and contain a distinct main frequency component in the output signal. The test calculates the correlation between the time varying amplitudes of the main frequency component and the higher harmonics. The bispectrum tests [10, 11] combine testing for Gaussianity and linearity. Peyton Jones and Billings [12] suggest identifying ARMAX-models before and after applying the same linear filter to the input and output signal. In the case of a linear system the identified system models should be identical.

Finally, the methods relying on the Hilbert Transform and the Nonlinear Prediction Error Test will be discussed in this paper. The remaining tests of Table 1 are common knowledge in engineering, so they do not need to be described here.

### 3. Hilbert Transform

The Hilbert Transform can be used to detect nonlinear system behaviour in measured frequency response functions (FRF) of technical systems. Examples are given in [14, 16], the initial suggestion was given by Vinh.

The FRF of linear systems, the so called transfer function  $G(s)$ , can be expressed by the ratio of two polynomials

$$G(s) = \frac{N(s)}{D(s)}, \quad s = i\omega \quad (1)$$

Table 1: Overview of nonlinearity tests

	Time Domain	Frequency Domain
Assumptions on Input / Output Measured only	<ul style="list-style-type: none"> <li>• Test for linearity against second-order Volterra expansion [Keenan/Tsay] [5, 6]</li> <li>• Surrogate Data Tests [Theiler et al.] [7]</li> <li>• Method of Internal Harmonics Cross-Correlation [Dimentberg/Sokolov/Haenisch] [8, 9]</li> </ul>	<ul style="list-style-type: none"> <li>• Bispectrum test [10, 11]</li> <li>• Total harmonic distortion  <math display="block">THD = \sqrt{\frac{a_2^2 + a_3^2 + \dots}{a_1^2 + a_2^2 + a_3^2 + \dots}}</math> </li> </ul>
Input Measured / Output Measured	<ul style="list-style-type: none"> <li>• Superposition of input signals: [1]  <math>u_1 \rightarrow y_1, u_2 \rightarrow y_2</math>  <math>\Rightarrow u_1 + u_2 \rightarrow y_1 + y_2</math></li> <li>• Identification of ARMAX-models before and after filtering [Peyton Jones/Billings] [12]</li> <li>• Nonlinear Prediction Error Test [Bruns/Popp] [13]</li> </ul>	<ul style="list-style-type: none"> <li>• Hilbert Transform [14]  <math>G = \mathcal{H}(\mathcal{G})</math></li> <li>• Linear spectral density [15]  <math> S_{uy} ^2 = S_{uu} * S_{yy}</math></li> </ul>

where  $i$  is the imaginary unit and  $\omega$  the angular frequency of excitation.

Assuming that all poles have negative real parts (stable system) it can be shown that an integral transform of  $G(s)$  exists in the right  $s$ -halfplane that maps  $G(s)$  onto itself. Because of the similarity of this transform to the Hilbert Transform the result is denoted  $H(s)$ . Considering the symmetry properties of transfer functions with real coefficients,  $H(s)$  can be split into the real and imaginary part

$$\begin{aligned} \operatorname{Re}\{H(i\omega_c)\} &= -\frac{2}{\pi} \text{P.V.} \int_0^{+\infty} \frac{\omega \operatorname{Im}\{G(i\omega)\}}{\omega^2 - \omega_c^2} d\omega \\ \operatorname{Im}\{H(i\omega_c)\} &= \frac{2\omega_c}{\pi} \text{P.V.} \int_0^{+\infty} \frac{\operatorname{Re}\{G(i\omega)\}}{\omega^2 - \omega_c^2} d\omega. \end{aligned} \quad (2)$$

Combining the left hand side of eqn. (2) results in a complex function that will be referred to as  $H(i\omega)$ . Any difference between  $H(i\omega)$  and the original  $G(i\omega)$  therefore indicates either a nonlinear or an unstable system.

Figure 1 shows three examples of applications of the Hilbert Transform. Figure 1a) shows the frequency response function  $G$  and the Hilbert Transform  $H$  of a simple mechanical oscillator with a softening spring. The characteristic deviation indicates nonlinear system behaviour. Figures 1b) and 1c) show the frequency response function of a cylindrical workpiece fixed in a grinding machine. The workpiece is excited tangentially to the grinding contact by an external force while the acceleration in the same direction is measured. In 1b) there is no contact between the workpiece and the grinding wheel. Subsequently, the measurement is repeated at a contact force of 120N (1c)). Clearly, nonlinearity due to friction and other effects is reflected in the difference between  $G$  and  $H$ . Comparing with Figure 1c), here also a softening spring characteristic can be assumed.

#### 4. Nonlinear Prediction Error Test

The Nonlinear Prediction Error Test [13, 17] is based on comparing prediction errors that are obtained by applying a linear and a nonlinear prediction model to the discrete time output signal  $y(k)$  of a system that has been subjected to a special input signal  $u(k)$ .

The input signal is composed of different random sections  $u_{Ri}(k)$ , that are required to excite the system, and repeated deterministic sections  $u_D(k)$ ,

$$u(k) = \{u_{R1}(k), u_D(k), u_{R2}(k), u_D(k), \dots\}. \quad (3)$$

In the consecutive  $u_D(k)$ -sections the system is excited by the same input signal, therefore, differences in the system

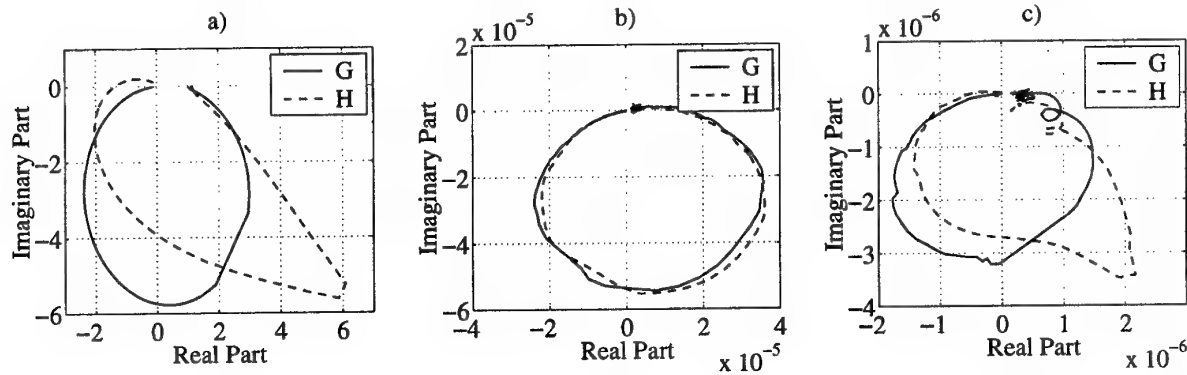


Figure 1: Nyquist Plots of frequency response functions ( $G$ ) and the corresponding Hilbert Transforms ( $H$ ): a) mechanical oscillator with a softening spring, b) response of a workpiece in a grinding machine, c) response of the same workpiece when it is pressed against the grinding wheel.

evolution with time are due to different initial conditions of the system states at the beginning of the  $u_D(k)$ -sections. Assuming the system under study is observable, different initial conditions are equivalent to different so called reconstructed state vectors

$$\mathbf{x}_r(k) = [y(k), y(k-1), \dots, y(k-d_E+1)], \quad (4)$$

that are composed of time lagged versions of the output signal  $y(k)$ . In the analysis of time series of nonlinear systems, the length of the reconstructed state vector is often called the embedding dimension  $d_E$ , this concept is also used here.

The nonlinear prediction model referred to in this test generates a prediction  $\tilde{y}(k+1)$  based on  $\mathbf{x}_r(k)$  using a local linear model (LLM). The model is fitted to the nearest neighbours of  $\mathbf{x}_r(k)$  found in the other  $u_D(k)$ -sections. The linear model used in this test is fitted to the input/output data of all  $u_D(k)$ -sections and is therefore referred to as the global linear model (GLM).

Figure 2 shows the results of applying the test to a simulated chain of three lightly damped nonlinear mechanical oscillators. The curves represent the mean of the absolute value of the prediction error  $e$  divided by the root mean square of the entire signal. The prediction error is plotted for both models, GLM and LLM versus the above mentioned embedding dimension used. The three plots are associated with different values of the scaling parameter  $\beta$  of the cubic component of the nonlinear restoring force

$$f_R(x) = (1 + \beta x^2)x. \quad (5)$$

In the case of the linear system ( $\beta = 0$ ) the two models yield equal prediction errors. The prediction error drops to a value near zero when the embedding dimension  $d_E$  reaches the number of system states that is equal to 6, corresponding to 3 mechanical degrees of freedom. When the system becomes nonlinear ( $\beta \neq 0$ ), the prediction performance of the nonlinear model (LLM) beats that of the linear model (GLM). The prediction error of the nonlinear model decreases still after the embedding dimension has reached the number of system states. This reflects the well known fact that with a nonlinear system of order  $N$  up to  $2N + 1$  output values may be required to characterize the system's state.

## 5. Summary

After a brief introduction and overview of nonlinearity tests that can be found in the literature, two tests were described and results given.

The test based on the Hilbert Transform was applied to experimental data from a grinding machine. It was shown that nonlinear effects in the contact zone between grinding wheel and workpiece are reflected in the Hilbert Transform results. Comparing the characteristic deviation of the Hilbert Transform from the originally measured frequency response function to numerical examples can help identifying the type of nonlinearity in the system.

The Nonlinear Prediction Error Test characterizes the strength of nonlinearity in the observed system dynamics by comparing the prediction performance of a linear and a nonlinear model. In addition, an estimate is obtained

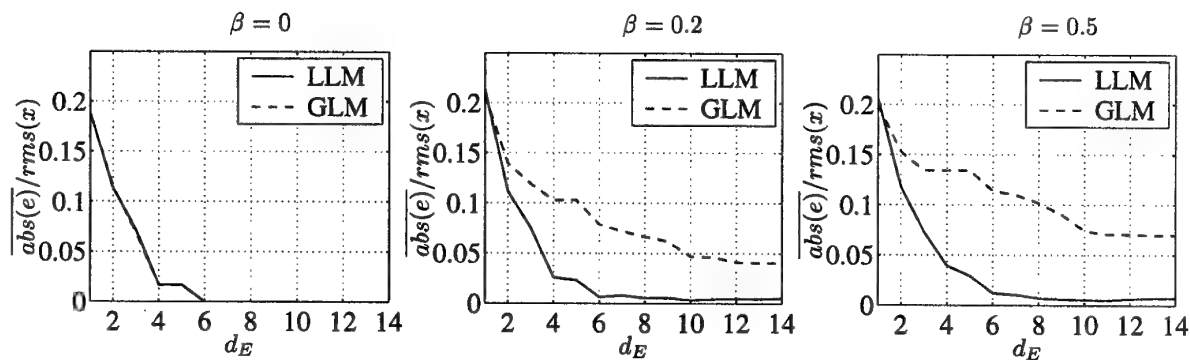


Figure 2: Results of the Nonlinear Prediction Error Test applied to a simulated chain of three mechanical oscillators. The system contains hardening springs due to cubic restoring forces scaled by the parameter  $\beta$ .

of the number of output values to be considered in an input/output model. A drawback of the method is that input-nonlinearities cannot be detected.

### Acknowledgements

The authors work is funded by the 'Bundesministerium für Bildung und Forschung'. The authors would like to thank Dipl.-Ing. O. Schütte for providing the grinding machine measurements.

### 6. References

- 1 HABER, R.: Nonlinearity tests for dynamic processes. IFAC Symp. on Identification and System Parameter Estimation, York, UK, 1985.
- 2 TONG, H.: Non-linear time series: a dynamical system approach. Oxford University Press, Oxford, 1990.
- 3 BARNETT, W. A. ET AL.: Robustness of nonlinearity and chaos tests to measurement error, inference method, and sample size. *Journal of economic behavior and organization*, **27** (1995), 301–320.
- 4 GABR, M. M.: Nonlinearity tests for bilinear systems. *Journal of computational and applied mathematics*, **40** (1992), 313–322.
- 5 KEENAN, D. M.: A Tukey nonadditivity-type test for time series nonlinearity. *Biometrika*, **72** (1985), 39–44.
- 6 TSAY, RUEY S.: Nonlinearity tests for time series. *Biometrika*, **73** (1986), 461–466.
- 7 THEILER, J., S. EUBANK, A. LONGTIN, B. GALDRIKIAN and J. D. FARMER: Testing for nonlinearity in time series: the method of surrogate data. *Physica / D*, **58** (1992).
- 8 DIMENTBERG, M. F. and A. A. SOKOLOV: Identification of restoring force non-linearity from a system's response to a white-noise excitation. *International Journal of Non-Linear Mechanics*, **26** (1991), 851–855.
- 9 DIMENTBERG, M. F. and H. G. HAENISCH: Method of Internal Harmonics Cross-Correlation for On-Line Detection of Structural Nonlinearities. Proceedings of the 12th Engineering Mechanics Conference, La Jolla, California, 1998.
- 10 SUBBA RAO, T. and M. M. GABR: A test for linearity of stationary time series. *J. Time Ser. Anal.*, **1** (1980), 145–158.
- 11 HINICH, M.: Testing for Gaussianity and linearity of a stationary time series. *J. Time Ser. Anal.*, **3** (1982), 169–176.
- 12 PEYTON JONES, J. C. and S. A. BILLINGS: Testing for nonlinearity using a prediction error filter. Research Report 332, Department of Control Engineering, University of Sheffield, 1988.
- 13 POPP, K. and J.-U. BRUNS: Nichtlinearitätstests und Bestimmung der Systemordnung. 4. Report to BMBF-Verbundproj.: Erkennen, Nutzen und Vermeiden nichtl. dyn. Effekte, Institut für Mechanik, Universität Hannover, 1999.
- 14 SIMON, M. and G. R. TOMLINSON: Use of the Hilbert transform in modal analysis of linear and non-linear structures. *Journal of Sound and Vibration*, **96** (1984), 421–436.
- 15 BENDAT, J. S. and A. G. PIERSON: Engineering applications of correlation and spectral analysis. John Wiley and Sons, New York, 1980.
- 16 BLOHM, P. W.: Untersuchung des nichtlinearen Übertragungsverhaltens von Strukturen mit lokalen Fügestellen. VDI Verlag GmbH, Düsseldorf, 1991.
- 17 POPP, K. and J.-U. BRUNS: Nichtlinearitätstests und Bestimmung der Systemordnung. 5. Report to BMBF-Verbundproj.: Erkennen, Nutzen und Vermeiden nichtl. dyn. Effekte, Institut für Mechanik, Universität Hannover, 2000.

Addresses: DIPL.-ING. JENS-UWE BRUNS, PROF. DR.-ING. KARL POPP,  
Universität Hannover, Institut für Mechanik, Appelstrasse 11, 30167 Hannover

GROTJAHN, M.; HEIMANN, B.

## Identification of industrial robots' nonlinear dynamics

*Simple linear joint controllers are still used in typical industrial robotic systems. The use of these controllers leads to non-negligible path deviations for complex tasks, e. g. laser-cutting, since nonlinear influences are neglected. Model-based compensation of nonlinear dynamics is the most common way to reduce these path deviations. For this, a parameter-linear multi-body model is generated and adapted to the robots' real behaviour by parameter identification methods. Different identification approaches are presented. The approaches are applied to the typical six d.o.f. industrial robot manutec-r15 and compared with respect to excitation problems and practical applicability. The application to path error prediction and compensation shows the quality of the identified model.*

### 1. Introduction

In recent years, robotic applications, which require not only position accuracy but also high path accuracy, have become more and more important not only in the academic but also in the industrial world. But, simple linear joint controllers are still used in typical industrial robot controls. Since nonlinear influences can not be compensated, the use of linear controllers leads to not acceptable path deviations.

Model-based compensation of nonlinearities is the most common way to reduce these errors. Since the dynamic parameters, like masses and moments of inertia, are usually not given by manufacturer and payload is not known in advance, the model's parameters typically have to be identified. Conventionally, linear parameter estimation techniques are used to solve this problem. For this, the equations of motion have to be derived in a form which is linear with respect to dynamic parameters.

In this article, two different approaches for identification of serial robots' rigid body dynamics are presented. The conventionally used identification scheme is compared to a so-called two-step approach with respect to industrial applicability and excitation problems. Identification is applied to the standard industrial robot *manutec-r15*. The identified model is used for simulation and reduction of tracking errors. The excellent results prove the quality of the model.

Although the gears have large impact on robot dynamics, the modeling of gear is not discussed in this article due to space restrictions. Normally, only the losses are taken into account by velocity dependent friction torques. Other effects like elasticity and backlash are neglected. For a detailed discussion of modeling and dynamic influence of friction see [1,2,4,9].

### 2. Rigid body model

The dynamic equation of the robot's rigid body model can be written as

$$Q = M(q)\ddot{q} + c(q, \dot{q}) + g(q) \iff Q = A(q, \dot{q}, \ddot{q})p. \quad (1)$$

The left equation represents the usual form of the dynamic equation with the mass matrix  $M(q)$  as well as the vectors of centrifugal and Coriolis forces  $c(q, \dot{q})$ , gravitational  $g(q)$  and joint torques  $Q$ . The right equation is the parameter linear form of the equation of motion. The base parameter vector  $p$  consists of the inertial and gravitational parameters of the links, e.g. masses and moments of inertia, and linear combinations of them. It has minimal order to guarantee identifiability of all elements  $p_i$  and can be derived automatically for any serial robot [7]. For typical industrial robots  $\dim(p)$  is relatively small because of the symmetric link structure. For the *manutec-r15* the dimension is 14.

The base parameters can be divided into two groups: The gravitational parameter vector  $p_g$  comprises the parameters that occur in  $g(q)$ , whereas the inertial parameters vector  $p_m$  consists of the parameter which only influence  $M(q)$  but not  $g(q)$ . This division leads to the following formulation of the equations of motion:

$$Q = \underbrace{A_{M,m}(q, \ddot{q})p_m + A_{M,g}(q, \ddot{q})p_g}_{M(q)\ddot{q}} + \underbrace{A_c(q, \dot{q})p}_{c(q, \dot{q})} + \underbrace{A_g(q)p_g}_{g(q)}. \quad (2)$$

### 3. Identification: Conventional approach

There exists a vast amount of literature on the identification of the rigid body model. However, most of the methods are variations of the same identification scheme. The robot is moved along a trajectory, where joint motion and torque are measured. Finally, the parameters are estimated by the use of the least-squares (LS) technique.

The LS method is based on the representation of robot dynamics shown on the right-hand side of equation (1). For a certain point in time  $T_i$ , measurements of  $m \leq n$  different axes  $\gamma_1 \dots \gamma_m$  are combined as

$$\mathbf{Q}_{T_i} = [Q_{\gamma_1, T_i} \dots Q_{\gamma_m, T_i}]^T \quad \text{and} \quad \Psi_{T_i} = [\mathbf{a}_{\gamma_1, T_i}^T \dots \mathbf{a}_{\gamma_m, T_i}^T]^T, \quad (3)$$

where  $Q_{\gamma_j}$  are the measured torques and  $\mathbf{a}_{\gamma_j}$  are the corresponding rows of  $\mathbf{A}$ . Further combination of measurements at  $r$  different time steps leads to the over-determined vector equation

$$\mathbf{Q} = \Psi \mathbf{p} + \mathbf{e}, \quad \mathbf{Q} = [\mathbf{Q}_{T_1}^T \dots \mathbf{Q}_{T_r}^T]^T \quad \text{and} \quad \Psi = [\Psi_{T_1}^T \dots \Psi_{T_r}^T]^T \quad (4)$$

with measurement vector  $\mathbf{Q}$ , observation matrix  $\Psi$ , parameter vector  $\mathbf{p}$  and the unknown error  $\mathbf{e}$ . Generally, equation (4) can be solved by using a pseudo-inverse of the observation matrix:

$$\hat{\mathbf{p}} = (\tilde{\Psi} \Psi)^{-1} \tilde{\Psi} \mathbf{Q}. \quad (5)$$

If  $\Psi$  is not a full rank matrix the pseudo-inverse does not exist. Therefore,  $\mathbf{p}$  must have minimal dimension as mentioned in the previous section. Minimizing the error in the LS sense leads to  $\tilde{\Psi} = \Psi^T$  and an estimated parameter vector

$$\hat{\mathbf{p}} = \arg \min_{\mathbf{p}} (\|\mathbf{e}\|) = (\Psi^T \Psi)^{-1} \Psi^T \mathbf{Q}. \quad (6)$$

This basic form of LS estimation can be refined by another choice of  $\tilde{\Psi}$  to e.g. total least-squares or instrumental variable method.

The main concern in using (5) and (6) is a proper choice of 'measurements' in order to ensure the excitation of all parameters. An insufficient excitation of one or more parameters would lead to a deterioration of the estimation result. An upper bound of the relative estimation error can be given by

$$\frac{\|\mathbf{p} - \hat{\mathbf{p}}\|}{\|\mathbf{p}\|} \leq \text{cond}(\Psi) \frac{\|\mathbf{e}\|}{\|\mathbf{Q}\|} \quad \text{with} \quad \text{cond}(\Psi) = \frac{\sigma_{\max}(\Psi)}{\sigma_{\min}(\Psi)}, \quad (7)$$

where  $\sigma_{\max}(\Psi)$  and  $\sigma_{\min}(\Psi)$  are the largest and the smallest singular value of  $\Psi$  [1,6]. The minimum of the condition of the observation matrix or the maximum of the smallest singular value are common criteria for optimizing trajectories in order to achieve excitation of all parameters. One problem is the computational burden if arbitrary trajectories are optimized. The computational costs can be reduced by using 7<sup>th</sup>-order polynomial trajectories [4]. Furthermore, such trajectories are shock- and jerkless such that unmodelled elasticities of joints are not excited by the driving torques.

The main problem of the optimization, however, is the fact, that standard industrial controls can only generate very simple trajectories. Therefore, the optimized trajectories can only be used in the field of robotics research, but usually not in industrial robotics.

### 4. Identification: Two-step approach

In contrast to the conventional identification approach, the two-step method requires only very simple trajectories and is, therefore, simply implementable in industrial robotics. It is based on the grouping in equation (2).

In the first step, gravitational torques and moments of inertia are 'measured' for a lot of different joint configurations, so-called operating points. Each measurement is carried out by moving one single axes 'back-and-forth' along some trapezoidal velocity profile in the neighborhood of the operating point. Such trapezoidal trajectories are very simple to generate by standard industrial controls. No specialized trajectories are needed as only one property has to be identified and the trajectories are chosen to excite just this property.

For measurements of gravitational torque, long periods with constant velocity have to be included. The mean between an averaged torque at forward and backward motion gives the desired gravitational torque. For the measurement of the moments of inertia motions with higher share of acceleration are used. Gravitation is compensated and the moment of inertia is identified in connection with a simple friction model. For more details on the measurements see [5,8].

The operating points are chosen by analyzing the structure of  $M(q)$  and  $g(q)$  to excite all dominant dependencies on the parameters  $p$  in order to include them in the estimation. Of course, also an optimization could be performed with respect to the common criteria (see section 3). A selection 'by hand', however, is sufficient as it leads to excellent results for the mentioned criteria. A further optimization would not yield recognizable improvements.

In the second step, the measurements of the gravitational torques are combined in

$$\underbrace{\begin{bmatrix} Q_{\gamma 1}^{(1)} \\ \vdots \\ Q_{\gamma m}^{(m)} \end{bmatrix}}_{\Gamma_g} = \underbrace{\begin{bmatrix} a_{g,\gamma 1}(q^{(1)}) \\ \vdots \\ a_{g,\gamma m}(q^{(m)}) \end{bmatrix}}_{\Psi_g} p_g + \underbrace{\begin{bmatrix} e^{(1)} \\ \vdots \\ e^{(m)} \end{bmatrix}}_{e_g}, \quad \gamma_i \in \{1 \dots n\}. \quad (8)$$

where  $Q_{\gamma i}$  are the measured gravitational torques,  $a_{g,\gamma m}(q^{(i)})$  are the corresponding rows of  $A_g$  for the given operating point  $q^{(i)}$  and  $e^{(i)}$  is an error that has to be introduced since  $p_g$  is not known. For the identification a weighted least squares criterion is introduced for  $e_g$  and leads to the estimation  $\hat{p}_g$

$$\hat{p}_g = \arg \min_{p_g} (e_g^T W e_g) \rightarrow \hat{p}_g = (\Psi_g^T W \Psi_g)^{-1} \Psi_g^T W \Gamma_g. \quad (9)$$

The diagonal weighting matrix takes into account the different ranges of the torque measurements by weighting them with respect to the maximum torque of the respective axis

$$W = \text{diag}([w_{\gamma 1} \dots w_{\gamma m}]), \text{ with } w_{\gamma i} = (Q_{\gamma i, \max})^{-1}. \quad (10)$$

The inertial parameter vector  $p_m$  is identified by combining the measurements of the moments of inertia  $M_{\delta i}^{(i)}$  of  $k$  different operating points. They are combined and the influence of the already known  $\hat{p}_g$  is compensated:

$$\underbrace{\begin{bmatrix} \hat{M}_{\delta 1}^{(1)} \\ \vdots \\ \hat{M}_{\delta k}^{(k)} \end{bmatrix}}_{\Gamma_M} - \underbrace{\begin{bmatrix} a_{M,g,\delta 1}(q^{(1)}, u_{\delta 1}) \\ \vdots \\ a_{M,g,\delta k}(q^{(k)}, u_{\delta k}) \end{bmatrix}}_{\Psi_M} \hat{p}_g = \underbrace{\begin{bmatrix} a_{M,\delta 1}(q^{(1)}, u_{\delta 1}) \\ \vdots \\ a_{M,\delta k}(q^{(k)}, u_{\delta k}) \end{bmatrix}}_{\Psi_M} p_m + \underbrace{\begin{bmatrix} e^{(1)} \\ \vdots \\ e^{(k)} \end{bmatrix}}_{e_M}. \quad (11)$$

Vectors  $u_{\delta i}$  stand for a vector of size of  $\ddot{q}$  with a 1 in the component of the measured inertia and zeroes elsewhere ( $\delta_i \in \{11 \dots nn\}$ ). An estimation for  $p_m$  is then found by

$$\hat{p}_m = \arg \min_{p_m} (e_M^T W e_M) \rightarrow \hat{p}_m = (\Psi_M^T W \Psi_M)^{-1} \Psi_M^T W \Gamma_M. \quad (12)$$

In the application to the *manutec-r15* both methods yield similar results. The disadvantage of the two-step approach is the higher measurement effort. In contrast to the conventional approach, which needs only one trajectory, a lot of experiments have to be performed. For the *manutec-r15* the procedure takes about 45 minutes. But on the other hand, no optimization is necessary as the choice of operating points could be performed 'by hand'. The single measurements are very simple. This makes the approach applicable to standard industrial robot systems which is the decisive advantage of the method.

## 5. Applications

One possible application of the identified robot model is the prediction of tracking errors [9]. Such a dynamic simulation could be integrated in off-line programming packages in order to notice path deviations in advance and so reduce teach-in costs. In figure 1 simulated and measured path deviations of the *manutec-r15* are shown as Euclidean distance ( $e_{ed} = \sqrt{e_x^2 + e_y^2 + e_z^2}$ ) for a vertical circle with 0.4m diameter. The identified model accurately predicts the dynamic behaviour of the robot.

Based on these results a compensation of the path deviations can be performed. One possibility is a model-based compensation by the well-known computed-torque method [3]. The necessary torques for a given trajectory  $q_d$  can be estimated using equation (1) and a previously identified friction model  $\hat{Q}_f(\dot{q})$ :

$$\hat{Q} = \hat{M}(q_d) \ddot{q}_d + \hat{c}(q_d, \dot{q}_d) + \hat{g}(q_d) + \hat{Q}_f(\dot{q}_d) = A(q_d, \dot{q}_d, \ddot{q}_d) \hat{p} + \hat{Q}_f(\dot{q}_d). \quad (13)$$

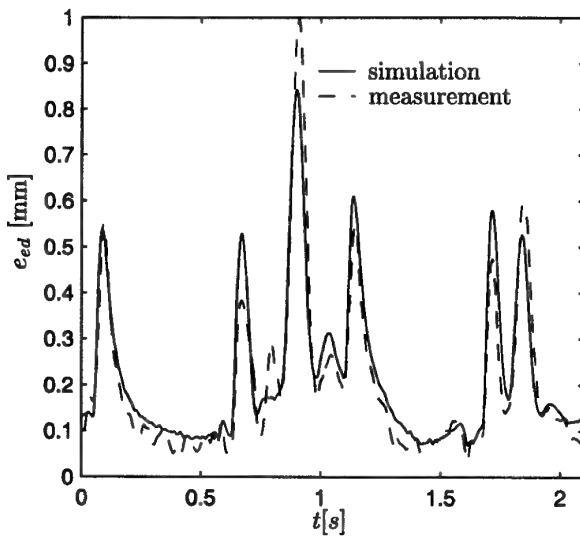


Fig. 1: Simulation and measurement of path errors for vertical circle.

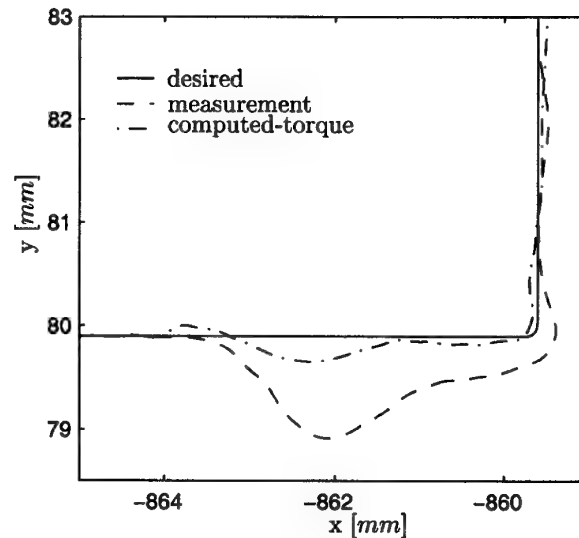


Fig. 2: Results of computed-torque method for an edge in x-y-plane.

Figure 2 shows that the computed-torque method yields impressive improvements of path accuracy. This is another proof for the quality of the identified model.

## 6. Summary

Two different approaches for rigid body identification of robots dynamics are presented and compared. The commonly used one leads to very little measurement costs but needs optimized trajectories which can usually not generated by standard industrial controls. This makes the application in industrial robotics difficult or even impossible. The two-step approach is much easier to implement since no optimized but only very simple trajectories are used. Therefore, it is applicable not only in research labs but also in industry.

Both method yield similar results. The model quality is shown by application to tracking error prediction and tracking error reduction by computed-torque method for the typical 6-d.o.f. industrial robot *manutec-r15*.

## 7. References

- 1 ARMSTRONG, B.: Dynamics for robot control: Friction modeling and ensuring excitation during parameter identification. PhD thesis, Stanford University, 1988.
- 2 ARMSTRONG-HÉLOUVRY, B.: Control of Machines with Friction. Kluwer Academic Publishers, Boston, 1991.
- 3 ASADA, H.; SLOTINE, J.-J. E.: Robot Analysis and Control. Wiley-Interscience Publication, 1986.
- 4 DAEMI, M.; HEIMANN, B.: Identification and compensation of gear friction for modeling of robots. In Proc. of the 11th CISM-IFTOMM Symp. on the Theory and Practice of Robots and Manipulators, 1996, 89–96.
- 5 DAEMI, M.; HEIMANN, B.: Separation of friction and rigid body identification for industrial robots. In Proc. of the 12th CISM-IFTOMM Symp. on the Theory and Practice of Robots and Manipulators, 1998, 35–42.
- 6 GAUTIER, M.; KHALIL, W.: On the identification of the inertial parameters of robots. In Proc. of the IEEE Conf. on Decision and Control, 1988, 2264–2269.
- 7 GAUTIER, M.; KHALIL, W.: Direct calculation of minimum set of inertial parameters of serial robots. IEEE Transactions on Robotics and Automation, 6(3), 1990, 368–373.
- 8 GROTJAHN, M.; DAEMI, M.; HEIMANN, B.: Friction and rigid body identification of robot dynamics. Int. Journal of Solids and Structures, 1999, accepted.
- 9 GROTJAHN, M.; HEIMANN, B.: Prediction of dynamic path errors for industrial robots. In Proc. of the 8th Int. Workshop on Robotics in Alpe-Adria-Danube Region, 1999, 77–82.

Addresses: DIPL.-ING. MARTIN GROTJAHN, PROF. DR.-ING. HABIL. BODO HEIMANN, Institut für Mechanik, Universität Hannover, Appelstraße 11, 30167 Hannover, Germany, email: grotjahn@ifm.uni-hannover.de



MERKWIRTH, C.; PARLITZ, U.

## Modeling chaotic and spatially extended systems

*Different aspects of local modeling of chaotic time series are discussed including cross validation methods and algorithms for fast nearest neighbor search. The resulting local models are used for predicting the temporal evolution of low-dimensional and spatio-temporal systems.*

### 1. Attractor reconstruction and local modeling

Measurements of dynamical systems typically provide discretely sampled scalar time series  $\{s^t\}$ ,  $t \in \mathbb{Z}$ . Using the method of delay embedding, states

$$\mathbf{x}^t = (s^t, s^{t-L}, \dots, s^{t-(D-1)L})$$

may be reconstructed from the data which are diffeomorphic images of the original states provided the *embedding dimension*  $D$  and the *delay time*  $L$  are chosen properly (Abarbanel, 1996; Kantz and Schreiber, 1997). This state space reconstruction may be used for deriving (black box) models describing the underlying process. In particular local descriptions of the dynamics in the reconstructed state space (typically around some reference point) in terms of *local models* (or: local predictors) turned out to be rather powerful. They are based on the assumption that neighboring states undergo similar evolutions and may thus be applied only when the flow  $\psi$  in the reconstructed state space is given by a continuous nonlinear function. Exploiting the special structure of delay embedding the approximation of the flow  $\psi$  can be reduced to the approximation of a scalar function  $f(\mathbf{x}) : \mathbb{R}^D \rightarrow \mathbb{R}$ ,  $\mathbf{x} \mapsto \tilde{s}$ :

$$\tilde{s}^{t+1} = f(\mathbf{x}^t) = f((s^t, s^{t-L}, \dots, s^{t-(D-1)L})). \quad (1)$$

The predicted time series value  $\tilde{s}^{t+1}$  can then be used for constructing the future state  $\tilde{\mathbf{x}}^{t+1}$  that enters the prediction of the next value  $\tilde{s}^{t+2}$  and so on. Of course, prediction errors accumulate with this kind of *iterative prediction*, but nevertheless iterated prediction is in general superior to direct prediction using a single, large prediction step.

For approximating the scalar function  $f(\mathbf{x})$  we shall use in the following local constant models that are based on preimage-image relations of the  $k$  nearest neighbors  $\mathbf{x}^{n_j}$ ,  $j = 1, \dots, k$  of a given state  $\mathbf{x}^t$ . Based on these relations two local constant prediction schemes have been proposed. *Absolute averaging*

$$\tilde{s}^{t+1} = \tilde{f}(\mathbf{x}^t) = \frac{1}{\sum_{j=1}^k w_j} \sum_{j=1}^k w_j s^{n_j+1} \quad (2)$$

where the result  $\tilde{s}^{t+1}$  of the prediction is a convex sum of the images  $s^{n_j+1}$  of the neighbors and thus lies inside the range spanned by the given data. This makes the prediction robust but also leads to poor results when extrapolation (or: generalization) is crucial for the prediction (for example, in sparsely occupied regions of the reconstruction space). In those cases *integrated averaging*

$$\tilde{s}^{t+1} = \tilde{f}(\mathbf{x}^t) = s^t + \frac{1}{\sum_{j=1}^k w_j} \sum_{j=1}^k w_j (s^{n_j+1} - s^{n_j}) \quad (3)$$

may lead to better results. To avoid discontinuities we use for both prediction schemes weights  $w_j = (1 - r_j^n)^n$  that depend on the relative distance  $r_j := \frac{d_j}{d_{k+1}}$  given by the distance  $d_j := d(\mathbf{x}^t, \mathbf{x}^{n_j})$  between the reference point  $\mathbf{x}^t$  and neighbor  $\mathbf{x}^{n_j}$  normalized by the distance to the  $(k+1)$ -th neighbor. For  $n > 0$  the resulting functions are  $(n-1)$ -fold continuously differentiable.

### 2. Leave-One-Out-Cross-Validation for local models

Cross validation (or: bootstrap) provides a simple but efficient way for choosing the free parameters of a given nonlinear model. Furthermore, this technique may be used to overcome the problem of overfitting where particular

features of the given data are incorporated into the model which are untypical for the dynamical process. For this purpose the available data are split into a *training set* which is used for deriving the model (parameters) and a *test set* that is used to check the quality of the model and to tune parameters. In order to improve statistical significance the division of the data into training and test set is repeated several times. An extreme version of cross validation is the so-called *Leave-One-Out-Cross-Validation* where a data set of length  $N$  is split into a training set of size  $N - 1$  and a test set of size 1. For global models this way of cross validation is extremely time consuming, because for each case the full global model has to be determined again. In the case of local models, however, just the single point which is left out has to be omitted in the selection of nearest neighbors and *Leave-One-Out-Cross-Validation* may thus be implemented very efficiently.

To evaluate the performance of the one-step-ahead-cross-validation we use the average mean square error

$$\text{MSE}_1^c = \frac{1}{|T_{\text{ref}}|} \sum_{t \in T_{\text{ref}}} \left( s^{t+1} - \tilde{f}_{t-c}^{t+c}(\mathbf{x}^t) \right)^2. \quad (4)$$

When using small sampling times one should avoid using neighbors in state space that lie on the same trajectory segment, because they contain only little information about the underlying deterministic dynamics. This is done here by omitting all neighbors with time indices  $nn_j \in [t - c, t + c]$  close to the time  $t$  of the current reference state  $\mathbf{x}^t$  as indicated with the notation  $\tilde{f}_{t-c}^{t+c}$ . In general, however, not the error of the one-step-ahead-cross-validation is most interesting, but the features of the models when applied iteratively. The normalized *multi-step-ahead-cross-validation error*

$$\text{NMSE}_{1,p}^c = \frac{N}{p|T_{\text{ref}}| \sum_{t=1}^N (s^t - \bar{s})^2} \sum_{t \in T_{\text{ref}}} \left( \left( s^{t+1} - \tilde{f}_{t-c}^{t+c}(\mathbf{x}^t) \right)^2 + \sum_{i=1}^{p-1} \left( s^{t+i+1} - \tilde{f}_{t+i-c}^{t+i+c}(\tilde{\mathbf{x}}^{t+i}) \right)^2 \right) \quad (5)$$

measures this performance over  $p$  iteration steps and is used in the following to quantify the quality of the derived models.

### 3. Fast nearest neighbors search

An important ingredient for local modeling are fast algorithms for searching the required nearest neighbors. The task of finding one or more nearest neighbors in a  $D_s$ -dimensional space can be used in many fields of data processing, e.g. information retrieval in database applications, data mining or, as in our case, for nonlinear time-series analysis where it may be used for modeling and prediction of time series, fast correlation sum computation (correlation dimension, generalized mutual information etc.), estimating the Renyi dimensions and Lyapunov exponents of experimental data or nonlinear noise reduction (Abarbanel, 1996; Kantz and Schreiber, 1997).

Nearest Neighbor searching and related problems of computational geometry have been extensively studied in the fields of computer science and pattern recognition and turned out not to fall into the class of computationally hard problems. Searching the nearest neighbor of all points in a data set of size  $N$  using a naive algorithm (which calculates the distances to every other point and picks out the smallest) is of order  $O(N^2)$ . However, for real world applications it would be useful to have an algorithm which is of order  $O(N \log(N))$ . A common approach to achieve this goal is to build up an auxiliary indexing data structure during a *preprocessing phase* which helps finding nearest neighbors during the *search phase*. Recently we proposed a new algorithm for searching nearest neighbors (Merkwirth et al., 2000) where during preprocessing, a hierarchical cluster tree is constructed which is then used in the search phase for efficiently locating neighboring points. The triangle inequality:  $d(x, z) \leq d(x, y) + d(y, z)$  is used in different ways to select for a given reference point (or: query point) the number of clusters that may contain possible candidates of nearest neighbors. Since the triangle inequality is valid in any metric, there is no limitation in what kind of metric is used to calculate distances.

### 4. Numerical examples

To illustrate the cross validation approach for local models time series of length  $N = 10000$  have been generated by numerical integration of different dynamical systems given in Table 1. For computing distances in reconstruction space we used a weighted metric that emphasises the most recent samples of the time series (McNames, 1998)

$$d_\lambda(\mathbf{x}^{t_1}, \mathbf{x}^{t_2}) = \sqrt{\sum_{i=0}^{D-1} \lambda^i (s^{t_1-iL} - s^{t_2-iL})^2}$$

	Dynamical System	Parameters	$T$	$\Delta T$	$D_I$
Baier-Sahle	$\dot{x}_1 = -x_2 + ax_1$ $\dot{x}_i = x_{i-1} - x_{i+1} \quad (i = 2, \dots, M-1)$ $\dot{x}_M = \epsilon + bx_M(x_{M-1} - d)$	$a = 0.28, \quad b = 4$ $d = 2, \quad \epsilon = 0.1$ $M = 5$	2000	0.2	4.26
Chua	$\dot{x}_1 = \alpha(x_2 - h(x_1))$ $\dot{x}_2 = x_1 - x_2 + x_3$ $\dot{x}_3 = -\beta x_2$ $h(y) = ym_2 + 0.5(m_2 - m_1)( y + c_0  -  y - c_0 )$	$\alpha = 9, \quad \beta = 14.286$ $m_1 = -\frac{1}{7}, \quad m_2 = \frac{2}{7}$ $c_0 = 1$	1000	0.1	2.30
Lorenz	$\dot{x}_1 = \sigma(x_1 - x_2)$ $\dot{x}_2 = rx_1 - x_2 - x_1x_3$ $\dot{x}_3 = x_1x_2 - bx_3$	$\sigma = -10$ $b = \frac{8}{3}$ $r = 28$	300	0.03	2.15
Rössler	$\dot{x}_1 = -x_2 - x_3$ $\dot{x}_2 = x_1 + ax_2$ $\dot{x}_3 = b + x_3(x_1 - c)$	$a = 0.45$ $b = 2$ $c = 4$	2000	0.2	1.97

Table 1: Dynamical systems used for generating chaotic time series. All time series are of length  $N = 10000$  and have been sampled with sampling time  $\Delta T$  after a transient time  $T$ . The last column contains the information dimensions  $D_I$  of the data sets.

System	$p$	$D$	$k$	$\lambda$	$w_n$	Modus	NMSE $_{1,p}^c$	NMSE $_{\max}/\text{NMSE}_{\min}$
Baier-Sahle	40	40	8	1.0	1	Integrated	0.050045	27.9
Chua	50	30	5	1.0	3	Integrated	0.053058	14.6
Lorenz	50	40	1	0.5	0	Integrated	0.084055	5.2
Rössler	80	40	1	0.7	0	Absolute	0.005964	20.7

Table 2: Results of cross validation of local models for chaotic time series generated with the dynamical systems given in Table 1. The last column shows the ratio of the errors of the best and the worst parameter combination.

where  $\mathbf{x}^{t_1}$  and  $\mathbf{x}^{t_2}$  are reconstructed state vectors and  $0 < \lambda \leq 1$  (including the ordinary euclidean metric for  $\lambda = 1$ ). The model selection consisted in a systematic computation of the normalized multi-step-ahead-cross-validation error (5) for all combinations of the following parameters and approximation methods

- embedding dimension  $D \in \{4, 8, 12, 16, 20, 25, 30, 40\}$
- exponent of the metric  $\lambda \in \{0.5, 0.6, 0.7, 0.8, 0.9, 1.0\}$
- number of neighbors  $k \in \{1, \dots, 8\}$
- type of weight function  $w_j = (1 - r_j^n)^n$  specified by the power  $n$
- absolute (2) or integrated averaging (3)

The parameter  $c$  equaled the mean return time, i.e., half of the value of the smallest time shift  $i$  for which the distance  $d_i = d(\mathbf{x}^r, \mathbf{x}^{r+i})$ ,  $i = 0, 1, \dots$  decreases again. For the delay time we used  $L = 1$ . This leads to rather high values of the optimal embedding dimension  $D$  but the used algorithm for finding nearest neighbors (Merkwirth et al., 2000) depends mainly on the (fixed) fractal dimension of the point set, while the embedding dimension has no strong influence on the computation time. Each time series was splitted 1000 times into training and test set. The resulting optimal parameter combinations are shown in Table 2. Figure 1 shows the results of iterative predictions using the optimal parameter values. The length of the time series shown equals  $4p$ , i.e. four times the prediction horizon  $p$  given in Table 2.

## 5. Spatially extended systems

Local modeling can also be applied to data from spatio-temporal systems. In this case, however, the dynamics is usually high-dimensional and special reconstruction techniques are necessary that exploit the fact that the data originate from an extended system. For systems whose dynamics is governed by local interactions in (configuration) space, the reconstruction of *local states* has turned out to be a useful method for analyzing and predicting the

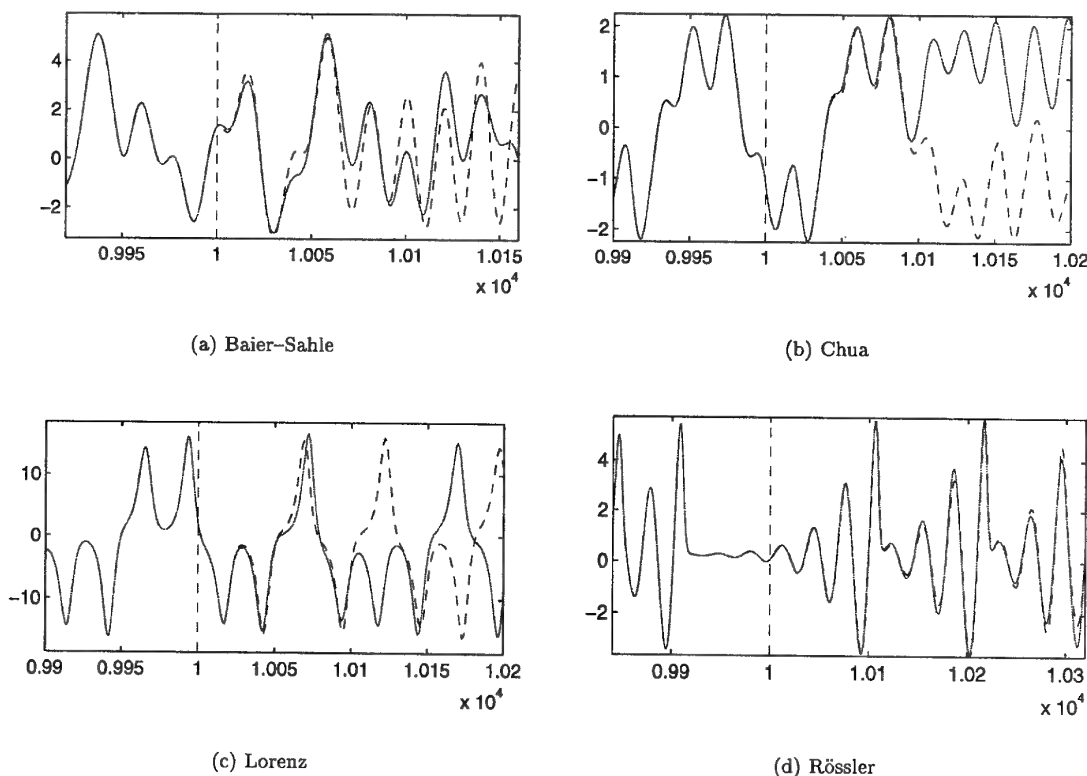


Figure 1: Iterated predictions computed with the optimal parameter combinations given in Table 2. The dashed curves show the true evolution and the solid lines give the prediction starting at the end of the training set as indicated by the vertical dashed line.

corresponding spatio-temporal time series (Parlitz and Merkwirth, 2000). With this reconstruction method samples from local space-time regions are used to create local states. Based on these states all methods described above can be used for predicting the underlying dynamics.

All numerical simulations presented in this paper were performed using Matlab programs from our nonlinear dynamics software package TSTOOL (<http://www.dpi.physik.uni-goettingen.de/tstool/>).

### Acknowledgements

*The authors acknowledge support by the Bundesministerium für Bildung und Forschung, grant no. 13N7038/9 and thank the members of the Nonlinear Dynamics Group at the "Drittes Physikalisches Institut" for stimulating discussions and support.*

### 6. References

- 1 ABARBANEL, H.D.I.: Analysis of Observed Chaotic Data; Springer Verlag, New-York/Berlin/Heidelberg, 1996.
- 2 KANTZ, H. AND SCHREIBER, TH.: Nonlinear Time Series Analysis; Cambridge Univ. Press, Cambridge 1997.
- 3 McNAMES, J.: A Nearest Trajectory Strategy for Time Series Prediction; Proc. of the INTERNATIONAL WORKSHOP ON ADVANCED BLACK-BOX TECHNIQUES FOR NONLINEAR MODELING (1998) 112-128.
- 4 MERKWIRTH, C., PARLITZ, U. AND LAUTERBORN, W.: Fast Exact and Approximate Nearest Neighbor Searching for Nonlinear Signal Processing; Phys. Rev. E **62**(2) (2000).
- 5 PARLITZ, U. AND MERKWIRTH, C.: Prediction of Spatiotemporal Time Series Based on Reconstructed Local States Phys. Rev. Lett. **84**(9) (2000) 1890-1893.

Addresses: CHRISTIAN MERKWIRTH, PD DR. ULRICH PARLITZ, Universität Göttingen, Drittes Physikalisches Institut, Bürgerstraße 42-44, D-37073 Göttingen, Germany.

P.C. MÜLLER

## Nonlinearity estimation by PI-observers: Theory and applications

### 1. Introduction

Dynamical systems are often influenced by troublesome nonlinear effects such as Coulomb friction, hysteresis or backlash. In [4] an indirect measuring technique of the actual values of these nonlinearities has been presented by defining an extended linear state observer which includes an integral feedback of the measurement error additional to the usual proportional feedback. This PI - observer yields estimates of the time behaviour of the state and of the nonlinearities as well. In the following this method of nonlinearity estimation is presented, its theory is touched, and its potentiality of applications is shown.

### 2. Method of nonlinearity estimation

The problem under consideration is described in the state space by

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{N}\mathbf{n}(\mathbf{x}(t), t) + \mathbf{B}\mathbf{u}(t), \quad \mathbf{y}(t) = \mathbf{C}\mathbf{x}(t) + \mathbf{D}\mathbf{u}(t) \quad (1)$$

where  $\mathbf{x}, \mathbf{u}, \mathbf{y}$  denote the  $n$ -dimensional state vector, the  $r$ -dimensional control vector of known inputs (e. g. control inputs or known excitations) and the  $m$ -dimensional measurement vector, respectively. The vector  $\mathbf{n}(\mathbf{x}, t)$  represents  $p$  more or less unknown functions which are generally nonlinear but which may be in special cases linear functions with unknown parameters or external disturbances depending only on time. The matrices  $\mathbf{A}, \mathbf{N}, \mathbf{B}, \mathbf{C}, \mathbf{D}$  are of related dimensions. To avoid redundant formulations the conditions  $\text{rk } \mathbf{N} = p, \quad \text{rk } \mathbf{B} = r, \quad \text{rk } [\mathbf{C}, \mathbf{D}] = r$  are assumed to be satisfied.

The problem of nonlinearity estimation consists in the construction of an estimate  $\hat{\mathbf{n}}(t)$  of the nonlinearities  $\mathbf{n}(\mathbf{x}, t)$  on the basis of the measurements  $\mathbf{y}, 0 \leq t$ . For this, it is assumed that the quantities  $\mathbf{A}, \mathbf{N}, \mathbf{B}, \mathbf{C}, \mathbf{D}$  and the input signals  $\mathbf{u}, 0 \leq t$ , are known.

According to [4, 5, 7] an extended linear state observer

$$\begin{bmatrix} \dot{\hat{\mathbf{x}}}(t) \\ \dot{\hat{\mathbf{v}}}(t) \end{bmatrix} = \begin{bmatrix} \mathbf{A} - \mathbf{L}_x\mathbf{C} & \mathbf{N} \\ -\mathbf{L}_v\mathbf{C} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{x}}(t) \\ \hat{\mathbf{v}}(t) \end{bmatrix} + \begin{bmatrix} \mathbf{B} \\ \mathbf{0} \end{bmatrix} \mathbf{u}(t) + \begin{bmatrix} \mathbf{L}_x \\ \mathbf{L}_v \end{bmatrix} (\mathbf{y}(t) - \mathbf{D}\mathbf{u}(t)) \quad (2)$$

is designed yielding the estimation of the state,  $\hat{\mathbf{x}}$ , and of the nonlinearities,  $\hat{\mathbf{n}}(t) = \hat{\mathbf{v}}(t)$ . Assuming the initial condition  $\hat{\mathbf{v}}(0) = \mathbf{0}$  then the combined proportional and integral feedback properties of the observer (2) are explicitly shown by

$$\hat{\mathbf{n}} = \hat{\mathbf{v}} = -\mathbf{L}_v \int (\mathbf{y} - \hat{\mathbf{y}}) dt, \quad \dot{\hat{\mathbf{x}}} = \mathbf{A}\hat{\mathbf{x}} + \mathbf{B}\mathbf{u} + \mathbf{L}_x(\mathbf{y} - \hat{\mathbf{y}}) + \mathbf{N}\mathbf{L}_v \int (\mathbf{y} - \hat{\mathbf{y}}) dt \quad (3)$$

where  $\hat{\mathbf{y}} - \mathbf{y}$  is the measurement estimation error with  $\hat{\mathbf{y}}(t) = \mathbf{C}\hat{\mathbf{x}}(t) + \mathbf{D}\mathbf{u}(t)$ . The choice of the observer gain matrices  $\mathbf{L}_x, \mathbf{L}_v$  can be realized that the observer (2) is asymptotically stable if and only if the extended system is detectable [5, 7]. Hints for the observer design and bounds for the estimation errors were presented in [7].

### 3. Applications

The proposed method of nonlinearity estimation may be applied to the identification of nonlinearity characteristics, to parameter identification, to fault detection or to nonlinearity compensation in closed-loop control systems. The PI-observer (2) yields an estimate (3) of the time behaviour of the nonlinearities. If one is interested in the characteristics of the nonlinearities then additional informations are required about the structural dependencies of  $\mathbf{n}(\mathbf{x}(t), t)$  on  $\mathbf{x}$  and  $t$ . If this a priori information is available, e.g. that a friction characteristic depends on a certain velocity which

is represented by one state variable  $x_l$ , then the reconstruction of the nonlinear characteristic is determined by

$$\hat{n}(\hat{x}, t) = \hat{v}(t)|_{t=\arg \hat{x}(t)} . \quad (4)$$

The problem of parameter estimation is easily illustrated in the special case of  $n(x, t) = Kx(t)$  where  $K$  is a matrix of unknown parameters. Then (3) results in  $\hat{v}(t) = \hat{K}\hat{x}(t)$ . In general it is not reasonable to evaluate this equation pointwise. But applying correlation methods a result is obtained for a sufficiently large  $T$ :

$$\hat{K} = \frac{1}{T} \int_0^T \hat{v}(t) \hat{x}^T(t) dt \cdot \left( \frac{1}{T} \int_0^T \hat{x}(t) \hat{x}^T(t) dt \right)^{-1} \quad (5)$$

The problem of fault detection in dynamical systems is an application of the identification methods mentioned before. In this case a fault is defined as an unknown nonlinearity which is zero in the faultless state and non-zero if a fault appears. Then a non-vanishing estimate (3) indicates the fault. Also sometimes a fault may be modelled by a change of a parameter which can be indicated by a recursive realization of the parameters estimation (5). The compensation of nonlinear effects within a closed-loop control system is based on the method of disturbance rejection [8]. If  $u$  represents control inputs then a feedback control  $u(t) = -K_x \hat{x}(t) - K_v \hat{v}(t)$  is used. The gain matrix  $K_x$  of the state feedback can be designed as usual but the gain matrix  $K_v$  of the nonlinearity compensation has to be calculated in a specific manner. In the special case if the matching condition is satisfied we have  $K = -M$  if  $N = BM$ . But in general  $K_v$  has to be determined by a more complicated system of linear equations, cf. [5,6].

The proposed method for estimation and compensation of nonlinearities by PI-observers has been successfully applied to a number of technical applications. Hasenjäger [2] used the method to compensate nonlinearities in the position control of parabolic antenna. Ackermann [1] compensated Coulomb friction in a highly accurate position control of elastic robots. Without this compensation steady-state inaccuracies or even limit cycles had appeared due to slip-stick friction. The good results have been verified in simulations as well as in experiments. Another application was the design of improved independent joint control of industrial robots estimating and compensating the coupling effects among the axes of the robot [3, 6]. The application of the PI-observer method to fault detection problems has been demonstrated in [9, 10]. Söffker et al. [10] detected cracks in turbo rotors. The observer based method allows an improved early detection of faults avoiding damages. More recently these results have been confirmed and generalized [9]. The method of estimation and compensation of nonlinearities has been successfully proved in the field of identification and reduction of the influence of parasitic effects. The main advantages are its simplicity because only standard algorithms of linear system theory are required and its robustness against modelling errors in the input channels of the nonlinearities.

#### 4. References

- 1 ACKERMANN, J.: Positionsregelung reibungsbehafteter elastischer Industrieroboter, VDI-Fortschr.-Ber. Reihe 8, Nr. 180, VDI-Verlag, Düsseldorf, 1989.
- 2 HASENJÄGER, E.: Digitale Zustandsregelung für Parabolantennen unter Berücksichtigung von Nichtlinearitäten, VDI-Fortschr.-Ber. Reihe 8, Nr. 87, VDI-Verlag, Düsseldorf, 1985.
- 3 HU, R.: Positionsregelung von Robotern mit der Methode der Nichtlinearitäten-Schätzung und -Kompensation, VDI-Fortschr.-Ber. Reihe 8, Nr. 827, VDI-Verlag, Düsseldorf 2000.
- 4 MÜLLER, P.C.: Indirect Measurement of Nonlinear Effects by State Observers, in: W. Schiehlen (Ed.): Nonlinear Dynamics in Engineering Systems, pp. 205-215, Springer, Berlin-Heidelberg 1990.
- 5 MÜLLER, P.C.: Schätzung und Kompensation von Nichtlinearitäten mit Störgrößenbeobachtern, in: Engell, S. (ed.): Entwurf nichtlinearer Regelungen, Oldenbourg, München - Wien, 1995, pp. 142-161.
- 6 MÜLLER, P.C.: Robuste Positionsregelung von Robotern mittels Nichtlinearitäten-Schätzung und -Kompensation, at -Automatisierungstechnik 48 (2000) pp. 289-295.
- 7 MÜLLER, P.C., BASPINAR, C.: Convergence of Nonlinearity Estimations by Linear Estimators, Z. Angew. Math. Mech. 80 (2000), Supplement 2, pp. S325-S326.
- 8 MÜLLER, P.C., LÜCKEL, J.: Zur Theorie der Störgrößenaufschaltung in linearen Mehrgrößenregelungssystemen, Regelungstechnik 25 (1977), pp. 54-59.
- 9 PARK, R.W., MÜLLER, P.C.: A Contribution to Crack Detection, Localization and Estimation of the Depth in a Turbo Rotor, Proc. 2nd Asian Control Conf., Vol. III, pp. 427-430, Seoul 1997.
- 10 SÖFFKER, D., BAJKOWSKI, J. AND MÜLLER, P.C.: Detection of Cracks in Turbo Rotors - A New Observer Based Method, ASME J. Dynamic Systems, Measurements and Control 115 (1993), pp. 518-524.

Address: Prof. Dr. P.C. Müller, Safety Control Engineering, University of Wuppertal, D-42097 Wuppertal, Germany.  
E-mail: mueller@srm.uni-wuppertal.de

TOMLINSON, G.R.

## Frequency Response Functions: Validity and Usefulness in Nonlinear System Identification

This paper considers the role of Higher Order Frequency Response Functions (HFRFs) for characterising nonlinear structures which can be represented by polynomial equations. By considering only the leading diagonals of these multi-dimensional HFRFs, simplifications are possible which allows practical procedures for experimentally measuring these, albeit they are approximations to the exact HFRFs due to the truncation of the series.

### 1. Ideal and Measured HFRFs

In order to describe the relationship between the ideal higher order FRFs and the measured higher order frequency response functions it is necessary to briefly describe the role of the Volterra series when the input is idealised (non-physical) harmonic function  $x(t) = X e^{j\omega t}$ . For a non-linear system which can be represented by a Volterra model, the output and input relationship becomes,

$$y(t) = X e^{j\omega t} \int_{-\infty}^{\infty} h_1(\tau_1) e^{-j\omega\tau_1} d\tau_1 + X^2 e^{j2\omega t} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h_2(\tau_1, \tau_2) e^{-j2\omega\tau_2} d\tau_1 d\tau_2 + \dots$$

$$+ X^n e^{jn\omega t} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} h_n(\tau_1, \dots, \tau_n) e^{-jn\omega\tau_n} d\tau_1 \dots d\tau_n \quad (1)$$

Noting that the terms inside the integral signs are multi-dimensional Fourier transforms, equation (1) becomes,

$$y(t) = H_1(j\omega) X e^{j\omega t} + H_2(j\omega, j\omega) X^2 e^{j2\omega t} + \dots + H_n(j\omega, \dots, j\omega) X^n e^{jn\omega t} \quad (2)$$

Equation (2) simply shows the relationship between the linear (one-dimensional) FRF function,  $H_1(j\omega)$  and the leading diagonal terms of the higher order FRFs  $H_n(j\omega)$  at the excitation frequency  $\omega$ , which are simply the higher harmonics at  $2\omega$ ,  $3\omega$  etc. The FRFs in equation (2) are defined in terms of the one-dimensional Fourier transforms of the input and output terms as,

$$H_1(j\omega) = \frac{Y(j\omega)}{X(j\omega)}$$

$$H_2(j\omega, j\omega) = \frac{2Y(j2\omega)}{X(j\omega)^2} \quad (3)$$

$$H_n(j\omega, \dots, j\omega) = \frac{2^{n-1} Y(jn\omega)}{X(j\omega)^n}$$

For a sine wave input (which is the practical case) we obtain the following Higher Order Transfer Functions:-

$$TF_1 = \frac{Y(j\omega)}{X(j\omega)} = H_1(j\omega) + \frac{3}{4} H_3(j\omega, j\omega, -j\omega) X(j\omega)^2 + \frac{n}{2^{n-1}} H_n(j\omega, \dots, -j\omega) X(j\omega)^{n-1} + \dots \quad n = 1, 3, 5, \dots \quad (4)$$

$$TF_n(jn\omega) = \frac{Y(jn\omega)}{X(j\omega)^n} = \frac{1}{2^{n-1}} H_n(j\omega, \dots, j\omega) + \frac{(n+2)}{n2^{n+2}} H_{n+2}(j\omega, \dots, -j\omega) X(j\omega)^2 + \dots \quad n = 1, 2, 3, \dots \quad (5)$$

The additional terms i.e. those above  $H_1(j\omega)$  are referred to as degenerative terms on the  $H_1(j\omega)$  function arising from the fact that the sine wave input excites the even and the odd higher order FRFs. If we take, as an example, what we normally measure in practice which is defined by equation (4) we see that to a first approximation,

$$TF_1(j\omega) = H_1(j\omega) + O(X(j\omega))^2 \quad (6)$$

and to a second approximation,

$$TF_1(j\omega) = H_1(j\omega) + \frac{3}{4} H_3(j\omega, j\omega, -j\omega) X(j\omega)^2 + O(X(j\omega))^4. \quad (7)$$

Equation (7) shows that the classical first order transfer function is equal to the classical first order FRF only when,  $H_3(j\omega, j\omega, j\omega)$ ,  $H_5(j\omega \dots j\omega)$ , etc are zero, i.e. when the system is linear. Further, it clearly shows that as the level of input excitation spectrum  $X(j\omega)$  is increased the significance of the higher order terms increases in relation to  $X(j\omega)^{n-1}$ . This can be physically interpreted as an increasing distortion of the measured transfer function (bending to the right or left) of systems with a hardening or softening stiffness non-linearity, as the amplitude of the input excitation increases. Indeed, the well known distortions apparent in the transfer function of a Duffing oscillator subject to sinusoidal excitation are readily explained by this analysis [3].

Equation (5) can be applied experimentally very simply to measure the approximate HFRFs, referred to as Higher Order Transfer Functions:

$$TF_n(jn\omega) = \frac{Y(jn\omega)}{X(j\omega)^n}$$

i.e.  $TF_1(j\omega) = \frac{Y(j\omega)}{X(j\omega)}$  is obtained by exciting the structure with a sine wave at a given frequency and measure the response and force input amplitude and phase at this frequency.

The second order transfer function,

$$TF_2(j2\omega) = \frac{Y(j2\omega)}{X(j\omega)^2}$$

is obtained by exciting with a sine wave at frequency  $\omega$  and measuring the response at the  $2\omega$  component, dividing by the square of the excitation force level at the forcing frequency to form  $TF_2(j2\omega)$ . This is repeated for higher order transfer functions as desired. Such transfer functions can offer considerable insight into the type of nonlinearities present in dynamic structural testing. Examples of the application of these ideas can be found in references 1, 2, 3, 4.

## 2. References

1. FRACHEBOURG, A: Identification of nonlinearities: application of the Volterra model to discrete MDOF systems, Proceedings of the Florence Modal Analysis Conference (1991) 93-100.
2. STORER, D M, TOMLINSON, G R: Parametric models of non-linear systems using multi-dimensional frequency response functions, Proceedings of the European Forum on Aeroelasticity and Structural Dynamics, Aachen (1989) 263-272.
3. STORER, D M: PhD thesis, University of Manchester, Dynamic analysis of nonlinear structures using higher order frequency response functions (1991).
4. STORER, D M, TOMLINSON, G R: Recent developments in the measurement and interpretation of higher order transfer functions from non-linear structures, J of Mechanical Systems and Signal Processing 7(2), 173-189 (1993).



UNBEHAUEN, H.

## Identifikation nichtlinearer Systeme aus regelungstechnischer Sicht

*Die Identifikation einer Regelstrecke liefert ein mathematisches Modell, das heute meist als Grundlage für den Entwurf einer modernen Regelung dient. Da die meisten technischen Regelstrecken nichtlinear sind und häufig eine Linearisierung in einem Arbeitspunkt nicht ausreicht, spielen Verfahren zur Identifikation nichtlinearer Modelle von Regelstrecken sowohl für die Regelung als auch Überwachung und Fehlererkennung eine zunehmend wichtige Rolle. Der Beitrag gibt eine Übersicht über die wichtigsten in der Regelungstechnik verwendeten nichtlinearen Modellformen, wobei zwischen nichtparametrischen, parametrischen und semiparametrischen Modellen unterschieden wird. Es werden Methoden zur Identifikation der das jeweilige Modell beschreibenden Kenngrößen erwähnt, und abschließend wird eine vergleichende Einschätzung dieser Verfahren gegeben.*

### 1. Einleitung und Problemstellung

Die Regelung hat die Aufgabe, die Regelgröße  $y(t)$  schnell auf den gewünschten festen oder variablen Sollwert  $w(t)$  zu bringen. Dabei ist eine generelle Voraussetzung für die Parametereinstellung oder die Synthese von Reglern, daß das dynamische Ein-/Ausgangsverhalten der Regelstrecke in Form eines *dynamischen Modells* bekannt ist. Ein solches Modell kann in vielfältiger Form aufgestellt werden, so z.B. als Differentialgleichung, Differenzengleichung, Übertragungsfunktion, Übertragungsmatrix, künstlich-neuronales Netzwerk oder linguistisches Fuzzy-Modell. Grundsätzlich bestehen zwei Möglichkeiten zur Herleitung eines solchen Modells. Die auf der Basis der die Regelstrecke beschreibenden physikalischen Gesetze hergeleiteten mathematischen Modelle sind meist sehr kompliziert und beruhen häufig auch auf verschiedenen Voraussetzungen, die teilweise nur ungenau zu erfüllen sind. Eine andere Möglichkeit der Modellbildung beruht auf der Messung der Ein- und Ausgangssignale  $u(t)$  und  $y(t)$  der betreffenden Regelstrecke. Diese Vorgehensweise wird auch als *experimentelle Systemidentifikation* bezeichnet [1]. Sie bildet die Grundlage für die nachfolgenden Betrachtungen.

Die meisten industriellen Regelstrecken weisen nichtlineares Verhalten auf. Dies ist entweder eine Folge der meist nichtlinearen Stellglieder und der häufig nichtlinearen Meßverfahren oder der nichtlinearen statischen Kennlinie, sofern die Regelstrecke über einen großen Arbeitsbereich betrieben wird und somit eine Linearisierung zu einem bestimmten Arbeitspunkt nicht mehr in Frage kommt. Daraus resultiert der Wunsch, das Verhalten der nichtlinearen Regelstrecke durch entsprechende nichtlineare Modelle zu beschreiben. Solche Modelle sind insbesondere erforderlich, wenn nichtlineare Regler eingesetzt werden.

### 2. Nichtlineare Modellstrukturen

Zahlreiche Verfahren wurden zur Modellierung nichtlinearer Regelstrecken vorgeschlagen [2 bis 6]. Gemäß [2] lassen sich diese Verfahren in 4 Gruppen unterteilen: (i) Nichtparametrische Modelle, (ii) Parametrische Modelle, (iii) Semiparametrische Modelle und (iv) Lineare Multimodelle.

Auf einige typische Vertreter aus diesen Modellgruppen soll nachfolgend kurz eingegangen werden.

#### 2.1. Nichtparametrische Modelle

Diese Modellklasse wird durch eine nichtparametrische Darstellung, z.B. in Tabellenform oder graphischen Formen (Zeitverläufe, Ortskurven, spektrale Darstellungen oder klassische Phasen-Ebene-Darstellung) charakterisiert. Ein typisches Beispiel ist die Volterra-Reihe [7]. Diese Modellbeschreibung für das Eingangs-/Ausgangsverhalten kann als eine Verallgemeinerung des für lineare Systeme bekannten Faltungsintegrals über das Produkt der Impulsantwort  $g(t)$  mit dem zeitverschobenen Eingangssignal  $u(t)$  angesehen werden. Die Systemidentifikation besteht hierbei darin, anhand der gemessenen Ein- und Ausgangsgrößen  $u(t)$  und  $y(t)$  die in den Integraltermen

$$\underbrace{\int \cdots \int}_{i\text{-mal}} g_i(\tau_1, \dots, \tau_i) \prod_{j=1}^i u(t - \tau_j) d\tau_1 \dots d\tau_i \quad i = 1, 2, \dots$$

der Volterra-Reihe auftretenden "Volterra-Kerne"  $g_i$  durch numerische Entfaltung punktweise zu bestimmen. Es ist leicht nachzuweisen, daß der hierfür erforderliche Rechenaufwand bereits für kleine Werte von  $i$  außerordentlich groß wird.

## 2.2. Parametrische Modelle

Bei dieser Modellklasse erfolgt die Darstellung bevorzugt in Form nichtlinearer Differenzen- oder Differentialgleichungen, deren Parameter speziell im kontinuierlichen Fall auch oft eine physikalische Bedeutung haben.

### 2.2.1. Diskrete parametrische Modelle

Als eine der allgemeinsten Modellformen dieser Gruppe ist das *Kolmogorow-Gabor-Modell* zu nennen [8,9], das auf der Basis eines Polynomansatzes beruht. Unter Einbeziehung eines Störmodells stellt dieses Modell eine nichtlineare ARMAX-Struktur dar und wird deshalb auch als NARMAX-Modell bezeichnet [10,11]. Dieses Modell ist linear in den Parametern, so daß diese relativ einfach mit einem LS-Verfahren geschätzt werden können. Zur Reduzierung der prinzipiell hohen Anzahl von Modelltermen stehen leistungsfähige, selbsttätig ablaufende Algorithmen zur Strukturbestimmung zur Verfügung [11 - 13].

Bei den beiden Modellstrukturen der *Wiener* und *Hammerstein*-Modelle wird davon ausgegangen, daß sich das dynamische Verhalten des Systems in einem linearen Teilmodell und das statische Verhalten in einem nichtlinearen Teilmodell darstellen läßt. Beide Teilmodelle werden in Reihe geschaltet, wobei das nichtlineare Teilmodell beim Hammerstein-Modell am Eingang und beim Wiener-Modell am Ausgang angeordnet wird. Beiden Modellen wird am Ausgang noch ein lineares Störmodell überlagert. Zahlreiche technische Regelstrecken können durch diese Modellstrukturen recht gut beschrieben werden. Hervorzuheben ist, daß die Struktur des Hammerstein-Modells linear in den Parametern ist, so daß die Parameter wiederum durch ein LS-Verfahren geschätzt werden können. Dies ist allerdings beim Wiener-Modell nicht der Fall. Hier müssen die Parameter durch ein Gradientenverfahren bestimmt werden [14]. Das *bilineare Modell* ist als Sonderfall bereits im Kolmogorow-Gabor-Modell enthalten. Die Nichtlinearität tritt hierbei nur in Form von Produkten zwischen Meßwerten der Ein- und Ausgangsgröße auf. Quadratische Terme und solche höherer Ordnung werden nicht berücksichtigt.

### 2.2.2. Kontinuierliche parametrische Modelle

Neben den kontinuierlichen Formen der bilinearen und der Hammerstein- und Wiener-Modelle lassen sich nichtlineare Systeme auch durch *integrierbare* Modelle der Form

$$\sum_{i=0}^{n_1} \sum_{j=0}^{n_2} a_{ij} \frac{d^i}{dt^i} [f_j[u(t), y(t)]] = 0$$

oder durch das *faltbare* Modell

$$\sum_{i=0}^{n_1} \sum_{j=0}^{n_2} \sum_{k=0}^{n_3} a_{ijk} g_k \left[ u(t), y(t) \frac{d^i}{dt^i} [f_j[u(t), y(t)]] \right] = 0$$

darstellen. Hierbei sind  $f_j$  und  $g_k$  bekannte Funktionen, während die Parameter  $a_{ij}$  bzw.  $a_{ijk}$  zu bestimmen sind. Zur Parameterschätzung werden diese Differentialgleichungen entweder durch Fourier- [15] oder zweckmäßiger Hartley-Modulationsfunktionen [16 - 18] in eine spektrale Darstellung gebracht. Da diese Darstellung linear in den Parametern wird, läßt sich die Parameterschätzung wiederum mit einem LS-Verfahren durchführen.

## 2.3. Semiparametrische Modelle

Diese neue Bezeichnung wurde eingeführt [5], um die Klasse der Modelle zu charakterisieren, die auf künstlich neuronalen Netzen (KNN) oder auf linguistischen Fuzzy-Regeln beruhen. Diese Modelle sind nicht streng in eine der zuvor behandelten Modellklassen einzuordnen. In beiden Fällen wird als Ziel der Identifikation eine gewisse Zahlenmenge geschätzt, die im Falle des KNN-Modells den Neuronengewichten und im Falle der Fuzzy-Modelle den Werten der Zugehörigkeitsfunktionen entspricht. Beide Klassen von Modellen sind auch in der Kombination von Neuro-Fuzzy-Modellen besonders gut zur Beschreibung von nichtlinearen Regelsystemen geeignet [19-21].

## 3. Kritische Bewertung der unterschiedlichen Modellstrukturen

Da die zuvor beschriebenen nichtlinearen Modellstrukturen sich teilweise ganz erheblich unterscheiden und für sehr unterschiedliche regelungstechnische Aufgabenstellungen entwickelt wurden, ist ein allgemeiner Vergleich hinsichtlich ihrer Leistungsfähigkeit nicht möglich. Auch sind manche Detailprobleme bis heute noch ungenügend geklärt.

Dennoch sollen nachfolgend einige wichtige Eigenschaften der unterschiedlichen Modellstrukturen angesprochen und für vier Modelle, als typische Vertreter der wichtigsten Modellklassen, dem

- Volterra-Modell (VM),
- Kolmogorov-Gabor (NARMAX)-Modell (KGM),
- kontinuierlichen faltbaren Differentialgleichungs-Modell (KFDM),
- künstlich neuronalen Netz-Modell (KNNM)

kritisch bewertet werden.

Da für eine solche Bewertung allgemeinverbindliche Maße bisher kaum zur Verfügung stehen, sollte diese Vorgehensweise nur als subjektive Betrachtung des Autors, allerdings auf der Grundlage umfangreicher Erfahrung mit diesen Modellstrukturen, gesehen werden. Eine wichtige Eigenschaft einer Identifikationsmethode ist ihre *Approximationsfähigkeit*. Diese ist ein unmittelbares Maß dafür, wie gut ein identifiziertes Modell in der Validierungsphase gemessene Signalverläufe der Systemausgangsgröße unter Vorgabe der zugehörigen gemessenen Eingangsgrößen approximiert. Ein wesentlicher Gesichtspunkt ist weiterhin der erforderliche *Rechenaufwand* für die Durchführung der Identifikation. Obwohl heute leistungsfähige Rechner zur Verfügung stehen, spielt der Rechenaufwand häufig die entscheidende Rolle, ob ein Verfahren im „on-line“ Betrieb, also auch in Realzeit in rekursiver Form eingesetzt werden kann. Dies ist ein wichtiger Gesichtspunkt bei der Realisierung adaptiver Regelsysteme. Weiterhin sind die Identifikationsverfahren unterschiedlich empfindlich gegen stochastische Störsignale. Leider sind die theoretischen Grundlagen für den Einfluß stochastischer Störungen in nichtlinearen Systemen bisher noch zu unbefriedigend erforscht, so daß weitgehend nur lineare Störmodelle dem Ausgangssignal überlagert werden. Daher können allgemeine Aussagen über die *Unempfindlichkeit* einer Identifikationsmethode gegenüber Rauschsignalen nicht gemacht werden. Hier ist man weitgehend auf Erfahrungswerte angewiesen.

Tabelle 1: Eigenschaften nichtlinearer Modelle

Eigenschaften	Modell	niedrig	mittel	groß
Approximationsfähigkeit	VM			x
	KGM			x
	KFDM			x
	KNNM			x
Rechenaufwand	VM			x
	KGM		x	
	KFDM	x		
	KNNM	x		
„on-line“ Fähigkeit	VM			nein
	KGM	(ja)		
	KFDM	ja		
	KNNM	ja		
Unempfindlichkeit gegen Rauschen	VM			(x)
	KGM			(x)
	KFDM			(x)
	KNNM			(x)
Konvergenz	VM		x	
	KGM			x
	KFDM			x
	KNNM			x
Datenvorverarbeitung	VM	erforderlich		
	KGM			nicht erforderl.
	KFDM	erforderlich		
	KNNM			nicht erforderl.

Bei den zur Systemidentifikation eingesetzten *Algorithmen* ist die *Konvergenz* zum „wahren“ Identifikationsergebnis, also den Schätzparametern und der Modellstruktur, von besonderer Wichtigkeit. Bei den Verfahren, deren Struktur linear in dem zu schätzenden Parameter ist, ist die *Identifizierbarkeit* mit einem LS-Verfahren gesichert, sofern das Ausgangssignal genügend erregt wird. Bei Gradientenverfahren ist dies nicht immer gewährleistet.

Schließlich sei noch darauf hingewiesen, daß bei manchen Verfahren eine *Vorverarbeitung der Meßdaten* erforderlich ist. Während bei den Verfahren zur Ermittlung diskreter parametrischer Modelle dies nicht erforderlich ist, empfiehlt

sich eine solche Datenvorverarbeitung häufig bei der Ermittlung kontinuierlicher parametrischer Modelle.

Tabelle 1 gibt eine Übersicht zur Bewertung der oben genannten vier Verfahren unter den hier genannten Gesichtspunkten.

#### 4. Zusammenfassung

Es stehen heute bereits zahlreiche bewährte Verfahren zur Identifikation nichtlinearer Systeme zur Verfügung. Im speziellen Anwendungsfall muß das geeignetste Modell sorgfältig ausgewählt werden, wobei man aufgrund der vorgegebenen Situation zuerst mit der einfachsten Modellstruktur beginnen sollte, und nur dann, falls diese nicht das gewünschte Ergebnis liefert, auf allgemeinere Modellklassen übergehen sollte. Häufig liefert auch eine genauere technisch-physikalische Analyse des zu modellierenden Systems Hinweise zur Wahl einer geeigneten nichtlinearen Modellstruktur.

In diesem kurzen Überblick wurde versucht, den augenblicklichen Entwicklungsstand zur Identifikation nichtlinearer Systeme aus regelungstechnischer Sicht darzustellen. Er ist für den interessierten Leser gedacht, um ihm einen leichten Einstieg in dieses wichtige Gebiet der dynamischen Systeme zu ermöglichen.

#### 5. Referenzen

- 1 UNBEHAUEN, H.: Regelungstechnik III, Vieweg-Verlag, Braunschweig, 1995
- 2 LJUNG, L.: System Identification – Theory for the User, Prentice-Hall, Englewood-Cliffs, N.J., 1999
- 3 SÖDERSTRÖM, T. UND STOICA, P.: System Identification, Prentice-Hall, Englewood Cliffs, N.J., 1989
- 4 UNBEHAUEN, H. UND RAO, G.P.: Identification of Continuous Systems, North Holland, Amsterdam, 1987
- 5 UNBEHAUEN, H.: Some New Trends in Identification and Modeling Nonlinear Dynamical Systems; Applied Mathematics and Computation **78** (1986), 279-297
- 6 BANYAS, C., HABER, R. UND KEVICZKY, L.: Some Estimation Methods for Nonlinear Discrete Time Identification; Proceedings, 3rd IFAC-Symposium on Identification and System Parameter Estimation, Delft, Niederlande (1973), 793-802
- 7 WIENER, N.: Nonlinear Problems in Random Theory; MIT Press, Cambridge, MA, 1958
- 8 GABOR, D., WILBY, W.P.L. UND WOODCOCK, R.: An Universal Nonlinear Filter, Simulator and Predictor which Optimizes itself by a Learning Process; IEE proceedings Part B. **108** (1961), 422-433
- 9 EYKHOFF, P.: System Identification, Wiley, London, 1974
- 10 LEONTARITIS, I. UND BILLINGS, S.: Input-Output Parametric Models for Nonlinear Systems (Part I and II); Int. J. Control **41** (1985), 303-344
- 11 KORTMANN, M. UND UNBEHAUEN, H.: Structure Detection in the Identification of Nonlinear Systems; APII (Automatique-Productive-Informatique-Industrielle) **22** (1988), 5-25
- 12 KORTMANN, M.: Die Identifikation nichtlinearer Ein- und Mehrgrößensysteme auf der Basis nichtlinearer Modellansätze; VDI-Fortschrittsbericht **177**, VDI-Verlag, Düsseldorf, 1989
- 13 HABER, R. UND UNBEHAUEN, H.: Structure Identification of Nonlinear Dynamic Systems – A Survey on Input-Output Approaches; Automatica **16** (1990), 651-677
- 14 KORTMANN, M. UND UNBEHAUEN, H.: Application of a Recursive Prediction Error Method to the Identification of Nonlinear Systems Using the Wiener Model; Proceedings, IMACS-Symposium on Modeling and Simulation for Control of Lumped and Distributed Parameter Systems, Villeneuve, France (1986), 281-285
- 15 PEARSON, A.E.: Explicit Parameter Identification for a Class of Nonlinear Input / Output Differential Operator Models; Proceedings, 31st IEEE Conference on Control and Decision, Tuscon, Arizona (1992), 3656-3660
- 16 PATRA, A. UND UNBEHAUEN, H.: Identification of a Class of Nonlinear Continuous-Time Systems Using Hartley Modulation Functions; Int. J. Control, **62** (1995), 1431-1451
- 17 DANIEL-BERHE, S. UND UNBEHAUEN, H.: Parameter Identification of Nonlinear Continuous-Time Systems by Hartley Modulation Functions Approach; J. of System Science **23** (1997), 5-22
- 18 DANIEL-BERHE, S.: Parameter Identification of Nonlinear Continuous-Time Systems Using the Hartley Modulation Functions Method, Cuvillier-Verlag, Göttingen, 1999
- 19 JUNGE, T. UND UNBEHAUEN, H.: Off-line Identification of Nonlinear Systems Using Structurally Adaptive Radial Basis Functions; Proceedings, 35th IEEE Conference on Decision and Control, Kobe, Japan (1996), 943-948
- 20 JUNGE, T.: Identifikation und lernende Regelung nichtlinearer Regelstrecken mittels neuronaler Netze; VDI-Fortschrittsbericht **807**, VDI-Verlag, Düsseldorf, 1977
- 21 KORTMANN, P.: Fuzzy-Modelle zur Systemidentifikation; VDI-Fortschrittsbericht **647**, VDI-Verlag, Düsseldorf, 1977

Adressen: PROF. DR. HEINZ UNBEHAUEN, LEHRSTUHL FÜR ELEKTRISCHE STEUERUNG UND REGELUNG, RUHR-UNIVERSITÄT BOCHUM, D-44780 BOCHUM,  
E-MAIL: UNBEHAUEN@ESR.RUHR-UNI-BOCHUM.DE

T. BÖHLKE, A. BERTRAM

## The 4th-Order Isotropic Tensor Function of a Symmetric 2nd-Order Tensor with Applications to Anisotropic Elasto-Plasticity

Dedicated to Prof. D. Gross on the event of his 60th birthday.

*The effective elastic properties of polycrystals can vary significantly with their crystallographic texture [7]. Since a correlation of elastic and plastic properties has been proven (see [8] and references therein), a phenomenological modeling of the crystallographic texture induced elastic anisotropy is of importance in the context of both elasticity and plasticity. In the present paper an evolution equation for the effective elasticity tensors of aggregates of cubic crystals is specified by means of the theory of isotropic tensor functions. It is shown that constraints forced by the elastic symmetry on the micro scale simplify the phenomenological equations significantly.*

### 1. Introduction

Notation: Linear mappings of 2nd-order tensors are written as  $\mathbf{A} = \mathbb{C}[\mathbf{B}]$ . The scalar product, the dyadic product, and the Euclidean norm are denoted by  $\mathbf{A} \cdot \mathbf{B}$ ,  $\mathbf{A} \otimes \mathbf{B}$ , and  $\|\mathbf{A}\| = (\mathbf{A} \cdot \mathbf{A})^{1/2}$ , respectively. *Lin* denotes the set of all 2nd-order tensors. *Sym* and *Orth* represent the sets of symmetric and proper orthogonal 2nd-order tensors.

Initially isotropic aggregates of crystalline grains show a texture-induced anisotropy of both their inelastic and elastic behavior when submitted to large inelastic deformations. The latter, however, is normally neglected, although experiments as well as numerical simulations clearly show a strong alteration of the elastic properties for certain materials. A source for such phenomena is a significant anisotropy of the corresponding physical property of the single crystals forming the aggregate. The main purpose of the present work is to derive the 4th-order isotropic tensor function of a symmetric 2nd-order tensor and to determine explicitly its irreducible part. This tensor function is necessary to formulate a phenomenological model for the evolution of the elastic properties polycrystals.

Generally, it is possible to decompose 4th-order elasticity tensors of arbitrary symmetry into a direct sum of orthogonal subspaces, on which the action of *Orth* is irreducible. The action of *Orth* on a vector space is said to be irreducible when there are no proper invariant subspaces. The harmonic decomposition has the form

$$\mathbb{C} = h_1 \mathbb{P}_1^I + h_2 \mathbb{P}_2^I + \mathbf{H}'_1 \otimes \mathbf{I} + \mathbf{I} \otimes \mathbf{H}'_1 + 4\mathbb{J}[\mathbf{H}'_2] + \mathbb{H}', \quad (1)$$

where

$$\mathbb{P}_1^I = \frac{1}{3} \mathbf{I} \otimes \mathbf{I}, \quad \mathbb{P}_2^I = \mathbb{I} - \mathbb{P}_1^I, \quad 4\mathbb{J}[\mathbf{A}] = (A_{im}\delta_{jn} + A_{in}\delta_{jm} + \delta_{im}A_{jn} + \delta_{in}A_{jm}) \mathbf{e}_i \otimes \mathbf{e}_j \otimes \mathbf{e}_m \otimes \mathbf{e}_n \quad (2)$$

[12, 13, 5, 3].  $\mathbf{I}$  denotes the 2nd-order identity tensor and  $\mathbb{I}$  represents the identity on symmetric 2nd-order tensors. The tensors  $\mathbf{H}'_1$ ,  $\mathbf{H}'_2$ , and  $\mathbb{H}'$  are irreducible, i.e. completely symmetric and traceless. A review concerning this representation is given in [6].  $h_1$  and  $h_2$  are called the first and second isotropic parts;  $\mathbf{H}'_1$  and  $\mathbf{H}'_2$  are the first and second deviatoric parts;  $\mathbb{H}'$  is the harmonic part. Irreducible 2nd-order tensors have five, and irreducible 4th-order tensors have nine independent components.

The symmetry group of  $\mathbb{C}$  is the intersection of the symmetry groups of its harmonic and deviatoric parts [6]. As a result, a cubic crystal symmetry forces  $\mathbf{H}'_1 = \gamma_1 \mathbf{I}$  and  $\mathbf{H}'_2 = \gamma_2 \mathbf{I}$ . From  $\text{tr}(\mathbf{H}'_1) = 0$  and  $\text{tr}(\mathbf{H}'_2) = 0$  one concludes  $\gamma_1 = 0$  and  $\gamma_2 = 0$ , respectively. Therefore, the tensors  $\mathbf{H}'_1$  and  $\mathbf{H}'_2$  vanish and the harmonic decomposition of the single crystal stiffness reduces to

$$\mathbb{C} = h_1 \mathbb{P}_1^I + h_2 \mathbb{P}_2^I + \mathbb{H}'. \quad (3)$$

Only in the case of a cubic crystal symmetry the deviatoric parts  $\mathbf{H}'_1$  and  $\mathbf{H}'_2$  vanish.

The effective elastic properties can be determined by orientational or volume averages of the local elasticity tensors. Examples are the arithmetic, the geometric, or the harmonic average [2]. In what follows we consider aggregates of

cubic crystals. The singlecrystalline grains are assumed to differ only by their crystallographic orientation. For the volume average of  $\mathbb{C}$  remains only

$$\bar{\mathbb{C}} = h_1 \mathbb{P}_1^f + h_2 \mathbb{P}_2^f + \bar{\mathbb{H}}^f. \quad (4)$$

Note that the volume average  $\bar{\mathbb{H}}^f$  of  $\mathbb{H}^f$  is irreducible. It is seen that a crystallographic texture evolution affects only the harmonic part of the stiffness. The same statement holds for the arithmetic mean [15] of local stiffnesses, the geometric mean [1, 9], and the harmonic mean [11] (see also [2]).

A simple phenomenological model for the texture induced elastic anisotropy is given by the following evolution equation

$$\frac{D}{Dt} \bar{\mathbb{H}}^f = \|\mathbf{D}_p'\| (\mathbb{G}'(\mathbf{N}_p') - d(I_p) \bar{\mathbb{H}}^f), \quad \mathbf{N}_p' = \frac{\mathbf{D}_p'}{\|\mathbf{D}_p'\|}, \quad I_p = \det(\mathbf{N}_p'), \quad (5)$$

where  $\mathbf{D}_p'$  is the macroscopic plastic strain-rate which is deviatoric. All quantities are formulated with respect to the (Lagrangian) undistorted configuration, which is invariant under changes of the observer.  $D(\cdot)/Dt$  denotes the material derivative. The plastic spin has not been taken into account in eqn (5). Furthermore, the driving term depends only on the direction  $\mathbf{N}_p'$  of  $\mathbf{D}_p'$ . The main problem is to find the general representation of the function  $\mathbb{G}'$ , which has to be irreducible. In section 2 the general 4th-order isotropic tensor function of a symmetric 2nd-order tensor is derived by means of the theory of isotropic tensor functions of 2nd-order tensors ([17, 4], see also [10, 16, 14]). In section 3 the corresponding irreducible part is determined. It is shown that the condition of irreducibility simplifies the representation considerably.

## 2. The 4th-Order Isotropic Tensor Function of a Symmetric 2nd-Order Tensor

In what follows, we derive the representation of a general, not necessarily polynomial, 4th-order isotropic tensor function  $\mathbb{G}$  of a symmetric 2nd-order tensor  $\mathbf{A} \in \text{Sym}$ . The tensor function  $\mathbb{G}$  is required to exhibit the index symmetries of elasticity tensors, i.e. the major symmetry and the symmetry in the first and second pair of indices

$$\mathbf{M} \cdot \mathbb{G}[\mathbf{N}] = \mathbf{N} \cdot \mathbb{G}[\mathbf{M}], \quad \mathbf{M} \cdot \mathbb{G}[\mathbf{N}] = \mathbf{M} \cdot \mathbb{G}[\mathbf{N}^T] = \mathbf{M}^T \cdot \mathbb{G}[\mathbf{N}] \quad \forall \mathbf{M}, \mathbf{N} \in \text{Lin}. \quad (6)$$

The starting point is the irreducible representation of a (symmetric) 2nd-order isotropic tensor function of two symmetric 2nd-order tensors [17, 4]

$$\mathbb{G}(\mathbf{A}, \mathbf{B}) = \sum_{\alpha=0}^7 g_{\alpha} \mathbb{G}_{\alpha}. \quad (7)$$

The eight symmetric tensor generators  $\mathbb{G}_{\alpha}$  are given by

$$\mathbf{I}, \mathbf{A}, \mathbf{A}^2, \mathbf{B}, \mathbf{AB} + \mathbf{BA}, \mathbf{A}^2\mathbf{B} + \mathbf{BA}^2, \mathbf{B}^2, \mathbf{AB}^2 + \mathbf{B}^2\mathbf{A}. \quad (8)$$

The  $g_{\alpha}$  are general functions of the 10 invariants of the functional basis of  $\mathbf{A}$  and  $\mathbf{B}$

$$\begin{aligned} &\text{tr}(\mathbf{A}), \text{tr}(\mathbf{A}^2), \text{tr}(\mathbf{A}^3), \text{tr}(\mathbf{B}), \text{tr}(\mathbf{B}^2), \text{tr}(\mathbf{B}^3), \\ &\text{tr}(\mathbf{AB}), \text{tr}(\mathbf{A}^2\mathbf{B}), \text{tr}(\mathbf{AB}^2), \text{tr}(\mathbf{A}^2\mathbf{B}^2). \end{aligned} \quad (9)$$

The representation (7) is called irreducible if the functional basis is irreducible and if none of the generators can be expressed as a linear combination of the other generators, formed with general functions  $g_{\alpha}$ . A functional basis is called irreducible if none of its elements can be expressed as a single-valued function of the other elements [4]. This definition of irreducibility differs from the one applied when discussing the index symmetries of the tensor  $\bar{\mathbb{H}}^f$ .

The function  $\mathbb{G}$  can be obtained by a linearization of  $\mathbb{G}$  in  $\mathbf{B}$

$$\mathbb{G}^{lin}(\mathbf{A}, \mathbf{B}) = \mathbb{G}(\mathbf{A})[\mathbf{B}]. \quad (10)$$

The linearized function  $\mathbb{G}^{lin}$  reads

$$\mathbb{G}^{lin}(\mathbf{A}, \mathbf{B}) = g_0^{lin} \mathbf{I} + g_1^{lin} \mathbf{A} + g_2^{lin} \mathbf{A}^2 + g_3^{lin} \mathbf{B} + g_4^{lin} (\mathbf{AB} + \mathbf{BA}) + g_5^{lin} (\mathbf{A}^2\mathbf{B} + \mathbf{BA}^2). \quad (11)$$

After the linearization, the scalar functions  $g_\alpha(\mathbf{A}, \mathbf{B})$  can be expressed in terms of  $\mathbf{B}$  and the new functions  $g_{ij}(\mathbf{A})$  that are isotropic in  $\mathbf{A}$

$$\begin{aligned} g_i^{lin} &= g_{i0}(\mathbf{A})\text{tr}(\mathbf{B}) + g_{i1}(\mathbf{A})\text{tr}(\mathbf{AB}) + g_{i2}(\mathbf{A})\text{tr}(\mathbf{A}^2\mathbf{B}), \quad (i = 0, 1, 2), \\ g_i^{lin} &= g_{i0}(\mathbf{A}), \quad (i = 3, 4, 5). \end{aligned} \quad (12)$$

A direct calculation yields the following representation

$$\begin{aligned} \mathbb{G}(\mathbf{A}) &= (3g_{00} + g_{30})\mathbb{P}_1^I + g_{30}\mathbb{P}_2^I + g_{11}\mathbf{A} \otimes \mathbf{A} + g_{22}\mathbf{A}^2 \otimes \mathbf{A}^2 + 2g_{40}\mathbb{J}[\mathbf{A}] + 2g_{50}\mathbb{J}[\mathbf{A}^2] \\ &+ g_{10}\mathbf{A} \otimes \mathbf{I} + g_{01}\mathbf{I} \otimes \mathbf{A} + g_{20}\mathbf{A}^2 \otimes \mathbf{I} + g_{02}\mathbf{I} \otimes \mathbf{A}^2 \\ &+ g_{21}\mathbf{A}^2 \otimes \mathbf{A} + g_{12}\mathbf{A} \otimes \mathbf{A}^2. \end{aligned} \quad (13)$$

From the requirement  $(6)_1$  one concludes

$$g_{10} = g_{01}, \quad g_{20} = g_{02}, \quad g_{12} = g_{21}. \quad (14)$$

As a result, the 4th-order isotropic tensor function  $\mathbb{G}$  reads

$$\begin{aligned} \mathbb{G}(\mathbf{A}) &= \sum_{\alpha=1}^9 G_\alpha(\mathbf{A})\mathbb{G}_\alpha(\mathbf{A}) \\ &= (3g_{00} + g_{30})\mathbb{P}_1^I + g_{30}\mathbb{P}_2^I + g_{11}\mathbf{A} \otimes \mathbf{A} + g_{22}\mathbf{A}^2 \otimes \mathbf{A}^2 + 2g_{40}\mathbb{J}[\mathbf{A}] + 2g_{50}\mathbb{J}[\mathbf{A}^2] \\ &+ g_{10}(\mathbf{A} \otimes \mathbf{I} + \mathbf{I} \otimes \mathbf{A}) + g_{20}(\mathbf{A}^2 \otimes \mathbf{I} + \mathbf{I} \otimes \mathbf{A}^2) \\ &+ g_{21}(\mathbf{A}^2 \otimes \mathbf{A} + \mathbf{A} \otimes \mathbf{A}^2). \end{aligned} \quad (15)$$

### 3. The Irreducible Part of the 4th-Order Isotropic Tensor Function of a Symmetric 2nd-Order Tensor

As mentioned before, an irreducible 4th-order tensor is symmetric and traceless with respect to every pair of indices. In this section we present the irreducible part  $\mathbb{G}'$  of the function  $\mathbb{G}$  (see (13)) by employing the procedure suggested by [5]. The irreducible part of a 4th-order tensor function  $\mathbb{G}$  is given by

$$\mathbb{G}' = \frac{1}{3}(\mathbb{G}) - \frac{1}{21}\{\hat{\mathbf{H}} \otimes \mathbf{I}\} + \frac{1}{105}\text{tr}(\hat{\mathbf{H}})(\mathbf{I} \otimes \mathbf{I}), \quad (16)$$

where

$$\hat{\mathbf{H}} = G_{iikl}\mathbf{e}_k \otimes \mathbf{e}_l + 2G_{ikil}\mathbf{e}_k \otimes \mathbf{e}_l. \quad (17)$$

$\{\mathbf{e}_k\}$  represents an orthonormal basis. The bracket formulae is defined by  $(\mathbf{A}, \mathbf{B} \in \text{Sym})$

$$\begin{aligned} \langle A_{ij}A_{kl} \rangle &= A_{ij}A_{kl} + A_{ik}A_{jl} + A_{il}A_{kj}, \\ \{A_{ij}B_{kl}\} &= A_{ij}B_{kl} + A_{ik}B_{jl} + A_{il}B_{kj} + B_{ij}A_{kl} + B_{ik}A_{jl} + B_{il}A_{kj}. \end{aligned} \quad (18)$$

Note, that  $\mathbb{G}$  has the major symmetric. Therefore,  $\langle G_{ijkl} \rangle = G_{ijkl} + G_{ikjl} + G_{iljk}$ . All components of  $\mathbb{G}'$  are linear functions of the components of  $\mathbb{G}$ .

Inspection of eqn (15) shows that only the following three of the nine 4th-order tensor generators contain non-vanishing irreducible parts

$$\mathbb{G}_3(\mathbf{A}) = \mathbf{A} \otimes \mathbf{A}, \quad \mathbb{G}_4(\mathbf{A}) = \mathbf{A}^2 \otimes \mathbf{A}^2, \quad \mathbb{G}_9(\mathbf{A}) = \mathbf{A}^2 \otimes \mathbf{A} + \mathbf{A} \otimes \mathbf{A}^2, \quad (19)$$

which are given by

$$\begin{aligned} \mathbb{G}'_3(\mathbf{A}) &= \frac{1}{3}(\mathbf{A} \otimes \mathbf{A}) - \frac{1}{21}(\text{tr}(\mathbf{A})\{\mathbf{A} \otimes \mathbf{I}\} + 2\{\mathbf{A}^2 \otimes \mathbf{I}\}) + \frac{1}{105}(\text{tr}(\mathbf{A})^2 + 2\text{tr}(\mathbf{A}^2))(\mathbf{I} \otimes \mathbf{I}), \\ \mathbb{G}'_4(\mathbf{A}) &= \mathbb{G}'_3(\mathbf{A}^2), \\ \mathbb{G}'_9(\mathbf{A}) &= \frac{1}{3}\{\mathbf{A}^2 \otimes \mathbf{A}\} - \frac{1}{21}(\text{tr}(\mathbf{A}^2)\{\mathbf{A} \otimes \mathbf{I}\} + \text{tr}(\mathbf{A})\{\mathbf{A}^2 \otimes \mathbf{I}\} + 4\{\mathbf{A}^3 \otimes \mathbf{I}\}) \\ &+ \frac{2}{105}(\text{tr}(\mathbf{A})\text{tr}(\mathbf{A}^2) + 2\text{tr}(\mathbf{A}^3))(\mathbf{I} \otimes \mathbf{I}). \end{aligned} \quad (20)$$

If the tensor  $\mathbf{A}$  is replaced by the direction  $\mathbf{N}' = \mathbf{A}'/\|\mathbf{A}'\|$  of its deviatoric part  $\mathbf{A}' = \mathbf{A} - \text{tr}(\mathbf{A})\mathbf{I}/3$ , then, the generators read

$$\mathbb{G}'_1(\mathbf{N}') = \frac{1}{3}\langle \mathbf{N}' \otimes \mathbf{N}' \rangle - \frac{2}{21}\{\mathbf{N}'^2 \otimes \mathbf{I}\} + \frac{2}{105}\text{tr}(\mathbf{N}'^2)\langle \mathbf{I} \otimes \mathbf{I} \rangle, \quad (21)$$

$$\mathbb{G}'_2(\mathbf{N}') = \frac{1}{3}\langle \mathbf{N}'^2 \otimes \mathbf{N}'^2 \rangle - \frac{1}{21}(\text{tr}(\mathbf{N}'^2)\{\mathbf{N}'^2 \otimes \mathbf{I}\} + 2\{\mathbf{N}'^4 \otimes \mathbf{I}\}) + \frac{1}{105}(\text{tr}(\mathbf{N}'^2)^2 + 2\text{tr}(\mathbf{N}'^4))\langle \mathbf{I} \otimes \mathbf{I} \rangle, \quad (22)$$

$$\mathbb{G}'_3(\mathbf{N}') = \frac{1}{3}\{\mathbf{N}'^2 \otimes \mathbf{N}'\} - \frac{1}{21}(\text{tr}(\mathbf{N}'^2)\{\mathbf{N}' \otimes \mathbf{I}\} + 4\{\mathbf{N}'^3 \otimes \mathbf{I}\}) + \frac{4}{105}\text{tr}(\mathbf{N}'^3)\langle \mathbf{I} \otimes \mathbf{I} \rangle. \quad (23)$$

Since  $\mathbf{N}'$  is traceless and normalized, the three functions  $G_1$ ,  $G_2$ , and  $G_3$  depend on the only non-constant principal invariant  $\det(\mathbf{N}')$  of  $\mathbf{N}'$ . As a result,  $\mathbb{G}'(\mathbf{N}')$  reads

$$\mathbb{G}'(\mathbf{N}') = G_3(I)\mathbb{G}'_3(\mathbf{N}') + G_4(I)\mathbb{G}'_4(\mathbf{N}') + G_9(I)\mathbb{G}'_9(\mathbf{N}'), \quad I = \det(\mathbf{N}'). \quad (24)$$

Within the presented evolution equation (5), the four scalar functions  $G_3(I_p)$ ,  $G_4(I_p)$ ,  $G_9(I_p)$ , and  $d(I_p)$  which depend on the scalar  $I_p = \det(\mathbf{N}'_p)$  remain to be identified.

#### 4. Conclusions

The theory of isotropy tensor functions of 2nd-order tensors is applied in order to formulate evolution equation of 4th-order elasticity tensors. The general representation theorem of a 4th-order isotropic tensor function of a symmetric tensor is derived. The irreducible part of this representation is determined explicitly. It is shown that the consideration of constraints given by the elastic symmetry on the micro scale simplifies the phenomenological equation significantly.

#### 5. References

- 1 K.S. ALEKSANDROV AND L.A. AISENBERG: A method of calculating the physical constants of polycrystalline materials. *Soviet Physics - Doklady*, **11** (1966) 323–325.
- 2 T. BÖHLKE AND A. BERTRAM: A minimum problem defining effective isotropic elastic properties. *Z. angew. Math. Mech.*, **80** (2000) S2:S419–S420.
- 3 J.P. BOEHLER, A.A. KIRILLOV, AND E.T. ONAT: On the polynomial invariants of the elasticity tensor. *J. Elast.*, **34** (1994) 97–110.
- 4 J.P. BOEHLER: Applications of Tensor Functions in Solid Mechanics. In *CISM Courses and Lectures*. Springer, Wien, 1987.
- 5 S.C. COWIN: Properties of the anisotropic elasticity tensor. *Q. J. Mech. appl. Math.*, **42** (1989) 249–266.
- 6 S. FORTE AND M. VIANELLO: Symmetry classes for elasticity tensors. *J. Elast.*, **43** (1996) 81–108.
- 7 U.F. KOCKS, C.N. TOME, AND H.R. WENK: *Texture and Anisotropy: Preferred Orientations in Polycrystals and Their Effect on Materials Properties*. Cambridge Univ. Pr., 1998.
- 8 C.-S. MAN: On the correlation of elastic and plastic anisotropy in sheet metals. *J. Elast.*, **39** (1995) 165–173.
- 9 S. MATTHIES AND M. HUMBERT: On the principle of a geometric mean of even-rank symmetric tensors for textured polycrystals. *J. Appl. Cryst.*, **28** (1995) 254–266.
- 10 A.C. PIPKIN AND A.S. WINEMAN: Material symmetry and restrictions on non-polynomial constitutive equations. *Arch. Rat. Mech. An.*, **12** (1963) 420–426.
- 11 A. REUSS: Berechnung der Fließgrenze von Mischkristallen auf Grund der Plastizitätsbedingung für Einkristalle. *Z. Angew. Math. Mech.*, **9** (1929) 49–58.
- 12 J.H. SCHOUTEN: *Der Ricci-Kalkül*. Springer, 1924.
- 13 A.J.M. SPENCER: A note on the decomposition of tensors into traceless symmetric tensors. *Int. J. Engng. Sci.*, **8** (1970) 475–481.
- 14 A.J.M. SPENCER: Theory of invariants. In C. Eringen, editor, *Continuum Physics*, pages 239–353. Academic Press, 1971.
- 15 W. VOIGT: *Lehrbuch der Kristallphysik*. Teubner Leipzig, 1928.
- 16 C.C. WANG: A new representation theorem for isotropic tensor functions, Part I and II. *Arch. Rat. Mech. An.*, **36** (1970) 166–223.
- 17 A.S. WINEMAN AND A.C. PIPKIN: Material symmetry and restrictions on constitutive equations. *Arch. Rat. Mech. An.*, **17** (1964) 184–214.

*Addresses:* T. BÖHLKE, A. BERTRAM, Otto-v.-Guericke-Universität Magdeburg, Institut für Mechanik, Postfach 4120, D-39016 Magdeburg



PAWELSKI, H.

## Theorie und Praxis des Dressierwalzens unter Berücksichtigung der Oberflächenveränderung

Beim Dressierwalzen insbesondere dünner Bänder kommen im Vergleich mit herkömmlichen Walzfällen verschiedene Besonderheiten erschwerend für die Modellierung hinzu: Elastische und plastische Deformationen des Bandes sind von gleicher Größenordnung (Verwendung des Prandtl-Reußschen Stoffgesetzes erforderlich), es tritt eine erhebliche Walzenabplattung auf (Ausbildung einer Förderzone mit nur eingeschränktem plastischen Fluß in der Mitte des Walzspalts), außerdem ist häufig die erzielte Dickenabnahme nicht wesentlich größer als die Rauhtiefen der Oberflächen von Walzen und Band (Oberflächenveränderung des Bandes muß in die Theorie mit einbezogen werden). Es wird ein Modell vorgestellt, das diese speziellen Gegebenheiten berücksichtigt, ohne jedoch zu große Rechenzeiten zu erfordern. Messungen an industriellen Dressiergeräten bestätigen die Verwendbarkeit des Modells.

### 1. Walzenabplattung

Die Vertikalverschiebung  $u$  der Walzenoberfläche durch elastische Verformung, die durch eine Vertikaldruckverteilung  $p(x)$  hervorgerufen wird, ist, siehe [3],

$$u(x) = \int_{-\infty}^{\infty} p(\xi) U(x - \xi) d\xi, \quad U(s) = -\frac{1 + \nu_W}{E_W \pi} \left\{ 1 + (1 - \nu_W) \ln \left[ \left( \frac{s}{R} \right)^2 \right] \right\}. \quad (1)$$

Die wesentlichen einen Walzfall kennzeichnenden Größen sind in der Tabelle auf der linken Seite von Abbildung 3 zusammengefaßt. Falls wir den Druck abschnittsweise als Polynom schreiben können, läßt sich eine geschlossene analytische Lösung für  $u(x)$  angeben. Sie setzt sich aus den folgenden Monomanteilen der Ordnung  $n$  zusammen:

$$u_n(x, x_1, x_2) := \int_{x_1}^{x_2} \xi^n U(x - \xi) d\xi = \frac{1 + \nu_W}{(n+1) E_W \pi} \left\{ x_1^{n+1} - x_2^{n+1} + (1 - \nu_W) \cdot \left[ 2 \sum_{j=1}^{n+1} \frac{x^{n-j+1}}{j} (x_2^j - x_1^j) + x_1^{n+1} \ln \frac{(x - x_1)^2}{R^2} - x_2^{n+1} \ln \frac{(x - x_2)^2}{R^2} - x^{n+1} \ln \frac{(x - x_1)^2}{(x - x_2)^2} \right] \right\}. \quad (2)$$

### 2. Einebnung der Oberfläche

Ein Bestandteil des Modells ist die Berücksichtigung der zunehmenden Kompression der Oberflächenschicht aufgrund des ansteigenden Drucks. Dazu benötigen wir einen Zusammenhang zwischen dem makroskopischen Druck  $p$  und dem Traganteil  $f$ , das ist der Quotient zwischen tragender Fläche der Rauheitsspitzen und der gesamten makroskopischen Kontaktfläche. Wir verwenden für trockene Reibung das Modell des Eindringens von Stempeln in einen Halbraum, vgl. auch [5]. Dabei ist  $k = k_f/\sqrt{3}$  die Schubfließspannung. Bei Schmierung, siehe [4], ist die Volumenkompression der Oberflächenschicht aufgrund der gefüllten Schmieraschen kleiner als bei trockener Reibung.

$$\frac{p}{2k} = \begin{cases} \left(1 + \frac{\pi}{2}\right) f, & \text{falls } f \leq \frac{1}{2}, \\ \frac{1 + \pi/2}{2\sqrt{1/f - 1}}, & \text{falls } f \geq \frac{1}{2}, \end{cases} \quad f = \begin{cases} \frac{p/(2k)}{1 + \pi/2}, & \text{falls } \frac{p}{2k} \leq \frac{1}{2} + \frac{\pi}{4}, \\ \frac{16[p/(2k)]^2}{16[p/(2k)]^2 + (2 + \pi)^2}, & \text{falls } \frac{p}{2k} \geq \frac{1}{2} + \frac{\pi}{4}. \end{cases} \quad (3)$$

### 3. Druck- und Dickenverteilung

Zur Aufstellung des Modells stellen wir geeignet zueinander passende, abschnittsweise aus Polynomen bestehende, Ansätze für den Vertikaldruck  $p(x)$  und die Gesamtdicke des Bandes  $h(x)$  auf, siehe Abbildung 1. Zur Ausbildung der Förderzone zwischen  $x_2$  und  $x_3$ , wie sie grundsätzlich auch vom Folienwalzen bekannt ist, siehe [1] und [2].

$$p_{01} = \frac{-p_1 x_0 + p_1 x}{x_1 - x_0}, \quad p_{12} = \frac{(p_1 x_2 - p_2 x_1) + (p_2 - p_1) x}{x_2 - x_1}, \quad p_{23} = p_{max} + c_1 x + c_2 x^2 + c_3 x^3 + c_4 x^4,$$

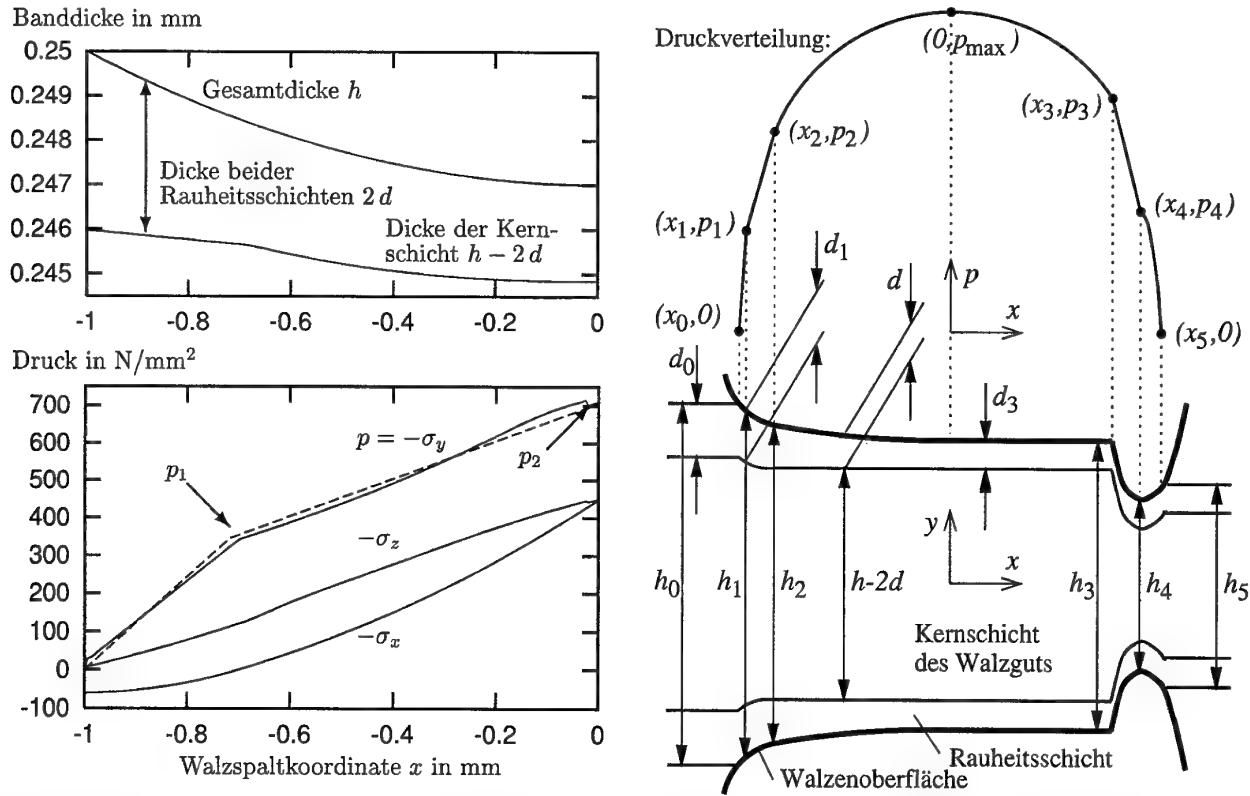


Abbildung 1: Links: Simulation des Druckanstiegs und der Kompression der Oberflächenschicht mittels Prandtl-Reuß-Streifentheorie, Vergleich mit linearisiertem Modell (gestrichelt).  $k_f=300 N/mm^2$ ,  $E=206000 N/mm^2$ ,  $\nu=0.3$ ,  $\sigma_0=\sigma_5=60 N/mm^2$ ,  $h_0=0.250 mm$ ,  $h_2=0.247 mm$ ,  $d_0=0.002 mm$ ,  $\mu=0.15$ . Rechts: Modell des Dressierwalzens.

$$p_{34} = \frac{(p_3 x_4 - p_4 x_3) + (p_4 - p_3) x}{x_4 - x_3}, \quad p_{45} = \frac{p_4 x_5 (x_5 - 2 x_4) + 2 p_4 x_4 x - p_4 x^2}{(x_5 - x_4)^2}. \quad (4)$$

Die Abkürzung  $p_{ij}$  bezieht sich auf  $p(x)$  im Bereich  $x_i \leq x \leq x_j$ . In  $p_{23}$  eliminieren wir  $c_2$  und  $c_4$

$$c_2 = [(p_3 - p_{max} - c_1 x_3 - c_3 x_3^3) x_2^4 + (p_{max} - p_2 + c_1 x_2 + c_3 x_2^3) x_3^4] / [x_2^2 x_3^2 (x_2^2 - x_3^2)],$$

$$c_4 = \{p_{max} (x_2^2 - x_3^2) - p_3 x_2^2 + x_3 [p_2 x_3 + x_2 (x_2 - x_3) (c_1 - c_3 x_2 x_3)]\} / [x_2^2 x_3^2 (x_2^2 - x_3^2)], \quad (5)$$

wobei  $p_{max}$ ,  $c_1$  und  $c_3$  freie Parameter bleiben. Wir wählen  $h(x)$  konsistent zu  $p(x)$  (die  $h_{ij}$  sind wie  $p_{ij}$  definiert):

$$h_{02} = c_{00} + c_{01} x + c_{02} x^2, \quad h_{23} = c_{20} + c_{22} x^2 \quad (x < 0), \quad h_{23} = c_{20} \quad (x \geq 0), \quad h_{35} = c_{30} + c_{31} x + c_{32} x^2,$$

$$\text{mit: } c_{00} = \{h_1 x_0 (x_0 - 2 x_2) + x_1 [h_0 (2 x_2 - x_1) + 2 c_{22} x_0 (x_1 - x_0) x_2]\} / q_1,$$

$$c_{01} = 2 [h_1 - h_0 + c_{22} (x_0^2 - x_1^2)] x_2 / q_1, \quad c_{02} = [h_0 - h_1 + 2 c_{22} (x_1 - x_0) x_2] / q_1,$$

$$c_{20} = [h_1 (x_0 - x_2)^2 - h_0 (x_1 - x_2)^2 - c_{22} (x_0 - x_1) x_2 (2 x_0 x_1 - x_0 x_2 - x_1 x_2)] / q_1,$$

$$c_{30} = \{(h_5 q_1 x_3 (x_3 - x_4) x_4 + [h_4 q_1 x_3 (x_5 - x_3) + q_2 x_4 (x_4 - x_5)] x_5\} / q_3,$$

$$c_{31} = \{q_1 [h_5 (x_4^2 - x_3^2) + h_4 (x_3^2 - x_5^2)] + q_2 (x_5^2 - x_4^2)\} / q_3,$$

$$c_{32} = \{q_1 [h_5 (x_3 - x_4) + h_4 (x_5 - x_3)] + q_2 (x_4 - x_5)\} / q_3,$$

$$q_1 = (x_0 - x_1) (x_0 + x_1 - 2 x_2), \quad q_3 = q_1 (x_3 - x_4) (x_3 - x_5) (x_4 - x_5),$$

$$q_2 = h_1 (x_0 - x_2)^2 - h_0 (x_1 - x_2)^2 - c_{22} (x_0 - x_1) x_2 (2 x_0 x_1 - x_0 x_2 - x_1 x_2). \quad (6)$$

Entscheidend ist die weitere Abnahme von  $h(x)$  aufgrund der Kompression der Oberflächenschicht bis zum Erreichen des Maximaldrucks  $p_{max}$ , obwohl die Dicke des Kerns  $h - 2d$  in der Förderzone ( $x_2 \leq x \leq x_3$ ) konstant bleibt.

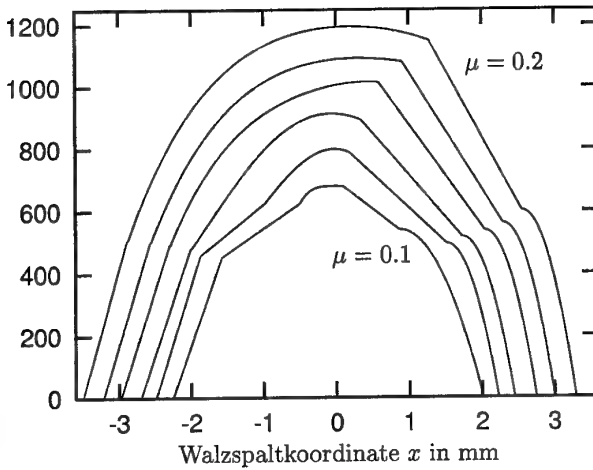
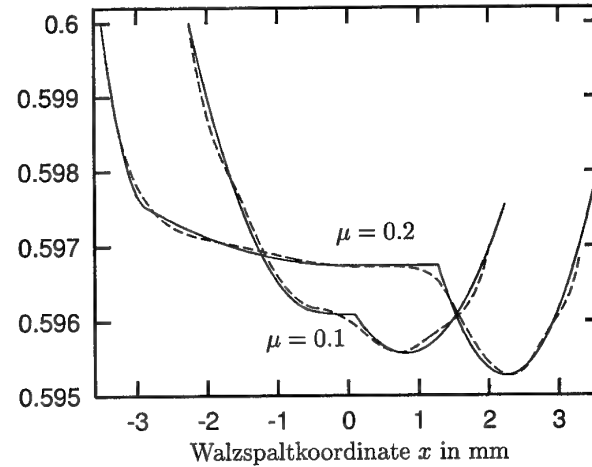
Vertikaldruck auf die Walze  $p$  in  $\text{N/mm}^2$ Dicke des Walzguts (mit Rauheitsschichten)  $h$  in mm

Abbildung 2: Beispiel.  $k_f=400 \text{ N/mm}^2$ ,  $E=E_W=206000 \text{ N/mm}^2$ ,  $\nu=\nu_W=0.3$ ,  $\sigma_0=\sigma_5=60 \text{ N/mm}^2$ ,  $R=300 \text{ mm}$ ,  $h_0=0.6 \text{ mm}$ ,  $h_5=0.597 \text{ mm}$ ,  $d_0=2 \mu\text{m}$ . Rechts ist  $h(x)$  (durchgezogen) mit  $h_W(x)$  (gestrichelt) verglichen.

#### 4. Prandtl-Reuß-Streifentheorie, linearisiertes Modell

Die Streifentheorie des Bandwalzens, [3], läßt sich bei Verwendung der inkrementellen Prandtl-Reuß-Gleichungen

$$d\varepsilon_x = \frac{1}{E} d\sigma_x - \frac{\nu}{E} d\sigma_y - \frac{\nu}{E} d\sigma_z + (2\sigma_x - \sigma_y - \sigma_z) \frac{d\varphi}{2k_f(\varphi)}, \quad (7)$$

$$d\varepsilon_y = \frac{1}{E} d\sigma_y - \frac{\nu}{E} d\sigma_z - \frac{\nu}{E} d\sigma_x + (2\sigma_y - \sigma_z - \sigma_x) \frac{d\varphi}{2k_f(\varphi)}, \quad (8)$$

$$0 = \frac{1}{E} d\sigma_z - \frac{\nu}{E} d\sigma_x - \frac{\nu}{E} d\sigma_y + (2\sigma_z - \sigma_x - \sigma_y) \frac{d\varphi}{2k_f(\varphi)} \quad (9)$$

und der differentiellen Form der von Misesschen Fließbedingung ( $\varphi$  Vergleichsformänderung)

$$(2\sigma_x - \sigma_y - \sigma_z)(2d\sigma_x - d\sigma_y - d\sigma_z) + (2\sigma_y - \sigma_z - \sigma_x)(2d\sigma_y - d\sigma_z - d\sigma_x) + (2\sigma_z - \sigma_x - \sigma_y)(2d\sigma_z - d\sigma_x - d\sigma_y) = 6k_f(\varphi)k'_f(\varphi)d\varphi \quad (10)$$

auf den elastisch-plastischen Fall erweitern, [6]. Zusätzlich wird die Kompression der Oberflächenschicht mittels  $d = d_0(1 - f + f^2/2)$  und Einsetzen von Gleichung (3) berücksichtigt. Ausgehend davon, siehe auch das Beispiel, Abbildung 1, links, wird folgendes linearisiertes Modell für das Materialverhalten im Ein- und Auslauf vorgeschlagen:

$$\begin{aligned} p_1 &= \frac{E}{1-\nu^2} \frac{(h_0 - 2d_0) - (h_1 - 2d_1)}{h_0}, \quad p_1 = \frac{2}{\sqrt{3}} k_f - \sigma_0 + \mu p_1 \frac{x_1 - x_0}{h_0}, \\ p_2 &= p_1 + \mu [p_1 + (1 - 2\mu) p_2] \frac{x_2 - x_1}{h_0}, \quad p_3 = p_4 + \mu [p_4 + (1 - 2\mu) p_3] \frac{x_4 - x_3}{h_0}, \\ p_4 &= \frac{E}{1-\nu^2} \frac{h_5 - h_4}{h_0}, \quad p_4 = \frac{2}{\sqrt{3}} k_f - \sigma_5 + \mu \frac{4p_4}{3} \frac{x_5 - x_4}{h_0}. \end{aligned} \quad (11)$$

Die Lösungen dieses Gleichungssystems lauten:

$$\begin{aligned} p_1 &= \frac{h_0 (2\sqrt{3} k_f - 3\sigma_0)}{3[h_0 - \mu(x_1 - x_0)]}, \quad p_2 = \frac{h_0 + \mu(x_2 - x_1)}{h_0 - \mu(1 - 2\mu)(x_2 - x_1)} p_1, \quad p_3 = \frac{h_0 + \mu(x_4 - x_3)}{h_0 - \mu(1 - 2\mu)(x_4 - x_3)} p_4, \\ p_4 &= \frac{h_0 (2\sqrt{3} k_f - 3\sigma_5)}{3[h_0 - 4\mu(x_5 - x_4)]}, \quad h_1 = h_0 - f_1(2 - f_1)d_0 - \frac{1-\nu^2}{E} h_0 p_1, \quad h_4 = h_5 - \frac{1-\nu^2}{E} h_0 p_4. \end{aligned} \quad (12)$$

Zur Berechnung des Traganteils  $f_1$  bei  $x = x_1$  nehmen wir an, daß der Druck bei Eintritt in die plastische Zone  $2k - \sigma_0$  ist, und setzen dies in die obere Zeile von Gleichung (3) ein:

$$f_1 = \frac{(2k - \sigma_0)/(2k)}{1 + \pi/2} = \frac{1 - (\sqrt{3}\sigma_0)/(2k_f)}{1 + \pi/2}. \quad (13)$$

$E_W, \nu_W$	Elastizitätsmodul, Poissonzahl der Walze
$R$	Radius der Arbeitswalze
$E, \nu$	Elastizitätsmodul, Poissonzahl des Bandes
$k_f$	Fließspannung (inklusive Abhängigkeit von $\dot{\varphi}$ )
$\sigma_0, \sigma_5$	Bandzugspannung am Ein-, Austritt
$h_0, h_5$	gesamte Banddicke am Ein-, Austritt
$d_0$	anfängliche Dicke einer Rauheitsschicht
$\mu$	Coulombsche Reibungszahl

Daten der Walzversuche:  $E = E_W = 206000 \text{ N/mm}^2$ ,  
 $\nu = \nu_W = 0.3$ ,  $R = 280 \text{ mm}$ ,  $k_f = 370 \dots 440 \text{ N/mm}^2$ ,  
 $\sigma_0, \sigma_5 = 50 \dots 200 \text{ N/mm}^2$ ,  
 $h_0 = 0.2 \dots 0.6 \text{ mm}$ ,  $d_0 = 4 \dots 5 \mu\text{m}$ ,  
 Verlängerung  $[(h_0 - 2d_0) - (h_5 - 2d_3)] / h_0 = 0.4 \dots 0.7\%$ ,  
 breitenbezogene Walzkräfte  $5 \dots 8 \text{ kN/mm}$ .

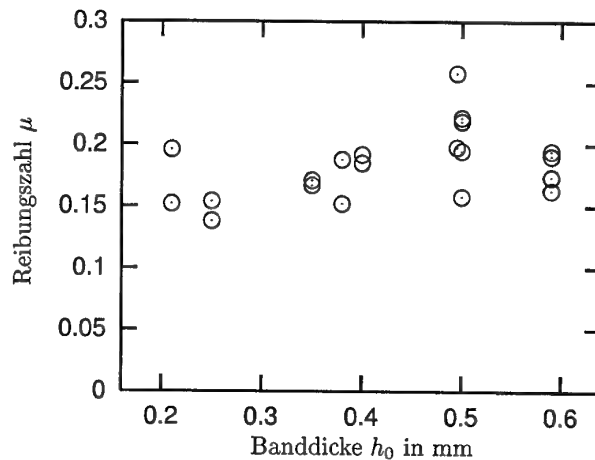


Abbildung 3: Tabelle auf der linken Seite: Einen bestimmten Walzfall kennzeichnende Größen. Rechte Seite: Aus gemessenen Walzkräften zurückgerechnete Reibungszahlen  $\mu$ .

Der Koeffizient  $c_{22}$  in den Gleichungen (6) ist schließlich (praktisch immer ist  $p_{max} > (1 + \pi/2)k \approx 1.48 k_f$  !):

$$c_{22} x_1^2 = 2(d_1 - d_3) = [f_3(2 - f_3) - f_1(2 - f_1)]d_0, \quad f_3 = 1 / \left[ 1 + \frac{1}{3} \left( \frac{1 + \pi/2}{p_{max}/k_f} \right)^2 \right]. \quad (14)$$

## 5. Algorithmus und Anwendungen

Ziel ist es nun, die noch unbekannten Größen  $x_0, \dots, x_5, p_{max}, c_1$  und  $c_3$  so zu bestimmen, daß die Walzspaltkontur  $h(x)$  mit der aus der zugehörigen Druckverteilung  $p(x)$  über die Walzenabplattung berechneten Kontur  $h_W(x)$  möglichst gut übereinstimmt:

$$\int_{x_0}^{x_5} [h(x) - h_W(x)]^2 dx \stackrel{!}{=} \text{Min}, \quad h_W(x) = 2 \left[ R - \sqrt{R^2 - x^2} + u(x) \right] + \text{const}. \quad (15)$$

Die Konstante wird dabei so gewählt, daß  $h_W(x_0) = h_0$ . Man beachte, daß alle Verteilungen, insbesondere auch die Abplattung bei Verwendung von (2), in analytischer Form vorliegen. Lediglich die Minimierung des Fehlerquadratintegrals (15) muß numerisch erfolgen, wobei sämtliche Ableitungen der Zielfunktion nach den Optimierungsparametern ebenfalls analytisch vorliegen.

Ein Beispiel für die Anwendung des Modells zeigt Abbildung 2. Das Modell wurde außerdem anhand der Prozeßdaten von Trockendressierstichen an Stahlbändern überprüft. In Abbildung 3, rechts, sind die aus den gemessenen Walzkräften zurückgerechneten Reibungszahlen zusammengestellt. Der eingehaltene Streubereich zeigt die Verwendbarkeit des Modells, wohingegen mit Standardwalztheorien (Walzenabplattung nach Hitchcock) keine sinnvollen Resultate zu erzielen sind.

## 6. Literatur

- 1 FLECK, N.A.; JOHNSON, K.L.: Towards a New Theory of Cold Rolling Thin Foil, International Journal of Mechanical Science, Vol. 29, No. 7, 1987, 507–524.
- 2 HARTUNG, H.G.: Die Statik und Kinematik des Folienwalzens, Dr.-Ing.-Dissertation, Umformtechnische Schriften 49, Verlag Stahleisen, Düsseldorf, 1994.
- 3 LIPPMANN, H.; MAHREHOLTZ, O.: Plastomechanik der Umformung metallischer Werkstoffe, Springer 1967.
- 4 PAWELSKI, H.: Evolution of surface roughness and corresponding friction during cold rolling, Advanced Technology of Plasticity, Proceedings of the 6th ICTP, Vol. 3, Springer 1999, 1931–1936.
- 5 PAWELSKI, H.: Modelling of temper rolling considering surface change of strip, Proceedings of the 8th Int. Conf. on Metal Forming in Krakow, A. A. Balkema 2000.
- 6 PAWELSKI, H.; PAWELSKI, O.: Technische Plastomechanik, Kompendium und Übungen. Verlag Stahleisen, Düsseldorf, 2000.

Adresse: DR.-ING. DIPL.-PHYS. HARTMUT PAWELSKI, SMS Demag AG, Zentralbereich Entwicklung, Umformtechnik, Postfach 23 02 29, 40088 Düsseldorf, Germany.

BROCKS, W.; BESSON, J.; CHABANET, O.; SCHEIDER, I.; STEGLICH, D.

## Modeling of Crack Growth in Sheet Metal

*Recent results of an investigation programme for the assessment of lightweight structures are reported. Special emphasis is laid upon the modeling aspect, and examples for the simulation of ductile crack growth by damage and cohesive zone models are given. The numerical simulations are compared with experimental data.*

### 1. Introduction

A realistic assessment of the residual strength of aircraft structures requires methods to characterize the crack growth resistance of the material and elastic-plastic analyses which are capable of simulating crack initiation and propagation. Conventional methods of fracture mechanics apply quantities like stress intensity factors,  $J$ -integral, energy release rate, or crack tip opening angle (CTOA), see e.g. [1, 2]. However, these integral measures can be applied to real components only under certain conditions, and the general lack of transferability of fracture data from specimens to structures impose important restrictions on the R-curve approach.

The so-called "local approach" to fracture provides a solution of the transferability problem by introducing local quantities and criteria for the degradation of the material. Especially, the application of micromechanically based models of the strength and toughness of materials with particles or inclusions has brought a better understanding of the mechanisms of ductile crack growth, see e.g. [3, 4]. These models have been successfully used to predict crack growth resistance curves for thick walled components of structural steels [5, 6]. Their application to thin walled components and high strength aluminum alloys, however, suffers from a number of specific problems:

- The stress state in sheet metal is totally different, namely the triaxiality is much lower and close to plane stress conditions whereas models of ductile damage have been established for large triaxialities.
- The fracture plane may shift from a normal to a 45° incline orientation to the applied load during crack growth, see figure 1.
- Rolled sheets generally show an anisotropic behaviour with respect to both, plastic hardening and void nucleating particles.
- Little is known about the specific damage phenomena in high strength aluminum in relation to its microstructure.

This requires enhanced constitutive models for deformation and damage especially adapted to the specific microstructure and properties of light weight materials and stress states in metal sheets, respectively.

### 2. Models of Ductile Damage and Failure

Ductile tearing of metals is dominated by the mechanisms of void nucleation at particles, void growth and coalescence. The "mesoscopic", i.e. averaged over the microstructure, inelastic deformation is described by a modified yield function and plastic potential,  $\Phi$ , including the "porosity" in terms of the void volume fraction,  $f$ , as an additional internal variable which is responsible for the "softening" of the material. The most common constitutive model for describing this process on a meso-scale was proposed by GURSON [7] and later modified by TVERGAARD and NEEDLEMAN [8, 9], hence addressed as GTN model:

$$\Phi = \left( \frac{\Sigma'_{ij} \Sigma'_{ij}}{R(\epsilon^{pl})} \right)^2 + f^* q_1 \cosh \left[ q_2 \frac{\Sigma_{kk}}{2 R(\epsilon^{pl})} \right] - (1 + q_3 f^{*2}) = 0 \quad (1)$$

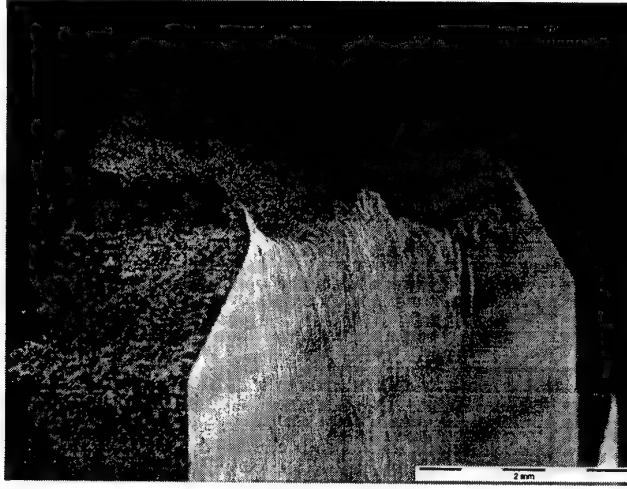


Figure 1: Ductile failure of a notched panel of Al 2024 showing slant fracture in the course of crack propagation

The GTN model has proven its ability to model crack initiation and growth in thin panels, see figure 2 showing the tunneling of the crack in the center plane. A simulated crack growth resistance curve of a center cracked panel of Al 2524, based on crack tip opening displacement, is shown in figure 3 (left), below. However, 3D modeling of large crack propagation is numerically costly. The present version does not yet account for any anisotropy of the plastic deformation and damage evolution. Furthermore, it cannot predict shear band formation over the specimen thickness and corresponding fracture in a  $45^\circ$  plane.

An alternative formulation of the plastic potential also introducing a scalar variable of porosity has been proposed by ROUSSELIER [10]:

$$\Phi = \frac{\sqrt{\frac{3}{2}\Sigma_{ij}\Sigma_{ij}}}{(1-f)R(\epsilon^{pl})} + \frac{\sigma_k}{R(\epsilon^{pl})} D f \exp\left[\frac{\Sigma_{kk}}{\sigma_k(1-f)}\right] - 1 = 0 \quad (2)$$

Commonly, this model is not supposed to yield major differences to the GTN model though despite the value of initial void volume fraction,  $f_0$ , all their other characteristic parameters are not comparable. Comparative analyses of round tensile bars and HILL-specimens however resulted in different predictions of the localisation of damage. The ROUSSELIER model turned out to have a better performance in predicting the formation of shear bands over the thickness of thin panels, see figure 2.

A phenomenological description of ductile tearing bases on BARENBLATT's idea of a "cohesive zone" at the crack tip where material separation is localized. The crucial problem in the application of a cohesive zone model (CZM) is, which decohesion law is appropriate for a given failure mechanism and, as in every model, how the material parameters can be identified. No possibilities of a direct measurement have been found yet. Several proposals have been made in the past, e.g. [11, 12, 13], which all base on a rather simple assumption, that the relation between traction,  $\Sigma_n$ , or shear,  $\Sigma_t$ , and the respective material separations,  $\delta_{cn}$ ,  $\delta_{ct}$ , is uniquely defined by two parameters per separation mode, namely the cohesive strengths,  $\sigma_{max}$ ,  $\tau_{max}$ , and the cohesive lengths,  $\delta_{cn}$ ,  $\delta_{ct}$ , or alternatively, the decohesion (or separation) energies,  $\Gamma_{nc}$ ,  $\Gamma_{tc}$ . ROSE [11] proposed a potential from which the normal and shear components are derived as

$$\Sigma_n = \sigma_{max} e \left\{ z \left( \frac{\delta_n}{\delta_{cn}} \right) - \frac{1}{2} z^2 \frac{\tau_{max}}{\sigma_{max}} \left( \frac{\delta_t}{\delta_{ct}} \right)^2 \right\} \exp \left[ -z \frac{\delta_n}{\delta_{cn}} \right]; \quad \Gamma_{nc} = \frac{9}{16} \sigma_{max} \delta_{cn} \quad (3)$$

and

$$\Sigma_t = \tau_{max} e \left\{ z \left( \frac{\delta_t}{\delta_{ct}} \right) \right\} \exp \left[ -z \frac{\delta_n}{\delta_{cn}} \right]; \quad \Gamma_{tc} = \sqrt{\frac{9}{32}} \tau_{max} \delta_{ct} \quad , \quad (4)$$

respectively, where  $e = \exp 1$  and  $z = 16e/9$ .

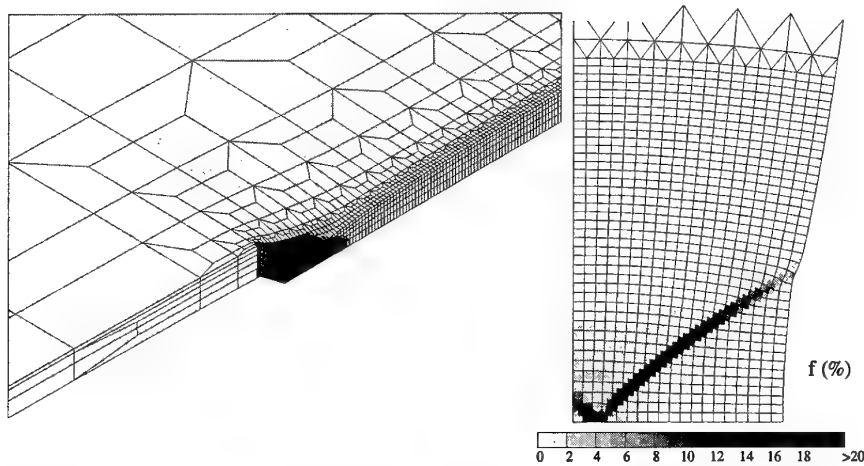


Figure 2: 3D finite element simulation of ligament necking and damage (dark zone) at the crack tip of a center notched panel of Al 2524 under tension showing tunneling of the crack front, FE code ABAQUS and GTN model (left); Formation of a shear band in a quarter model of a HILL-specimen, plane strain analysis with FE code ZEBULON and ROUSSELIER model (right)

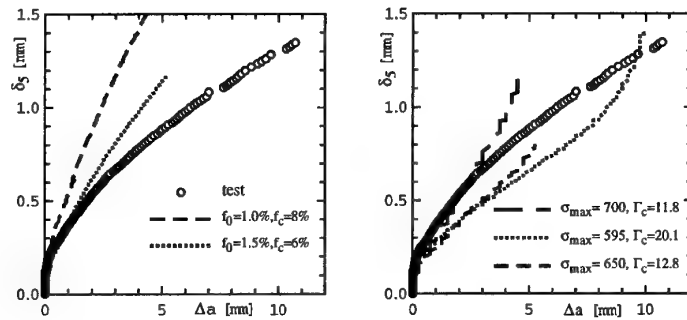


Figure 3: Simulation of CTOD based crack growth resistance of a center cracked panel of Al 2524 by GTN model (left) and CZM (right), FE code ABAQUS.

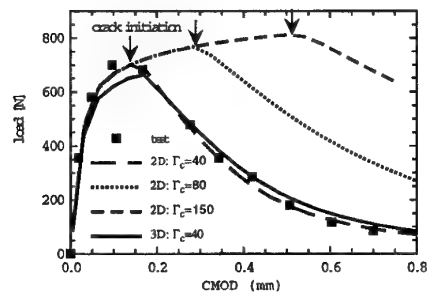


Figure 4: Simulation of deformation and crack growth in a laser welded micro SE(B) specimen by a CZM, influence of local separation energy,  $\Gamma_c$ , FE code ABAQUS.

This formulation has been used by NEEDLEMAN and other authors [12, 14]. The CZM relates microscopic to macroscopic parameters [15] and has also been successfully applied to simulate ductile crack growth in aluminum panels [16, 17], see figure 3 (right). It is numerically effective for large amounts of crack growth, it is applicable to characterize bimaterial interfaces and welds, see figure 4, and it even can be used to model combined normal and shear fracture.

### 3. References

- 1 HARRIS, C.E., NEWMAN, J.C., PIASCIK, R.S., AND STARNES, J.H. (1998): "Analytical Methodology for Predicting Widespread Fatigue Damage Onset in Fuselage Structures", *Journal of Pressure Vessel Technology*, Vol. 104, pp. 299-307.
- 2 CHEN, C.S., WAWRZYNEK, P.A., AND INGRAFFEA, A.R. (1999): "Crack Growth Simulation and Residual Strength Prediction in Airplane Fuselages", Report NASA/CR-1999-209115.
- 3 NEEDLEMAN, A. AND TVERGAARD, V. (1987): "An Analysis of Ductile Rupture at a Crack Tip", *J. of Mechanics and Physics of Solids*, Vol. 35, pp. 151-183.
- 4 STEGLICH, D. AND BROCKS, W. (1998): "Micromechanical Modelling of Damage and Fracture of Ductile Materials", *Fatigue and Fracture of Engineering Materials and Structures*, Vol. 21, pp. 1175-1188.
- 5 SUN, D.-Z., SIEGELE, D., VOSS, B. AND SCHMITT, W. (1988): "Application of Local Damage Models to the Numerical Analysis of Ductile Rupture", *Fatigue and Fracture of Engineering Materials and Structures*, Vol. 12, pp. 201-212.
- 6 BROCKS, W., KLINGBEIL, D., KNECKE, G., AND SUN, D.-Z. (1995): "Application of the Gurson Model to Ductile Tearing Resistance", *Second Symposium on Constraint Effects*, ASTM STP 1224, M. Kirk and A. Bakker, Eds., pp. 232-252.
- 7 GURSON, A. L. (1977): "Continuum Theory of Ductile Rupture by Void Nucleation and Growth: Part I - Yield Criteria and Flow Rules for Porous Ductile Media", *J. Engng. Materials and Technology* Vol. 99, pp 2-15.
- 8 TVERGAARD, V. (1982): "On Localization in Ductile Materials Containing Spherical Voids", *Int. J. Fracture* Vol. 18, pp. 237-252.
- 9 NEEDLEMAN, A. AND TVERGAARD, V. (1984): "An Analysis of Ductile Rupture in Notched Bars", *J. of Mechanics and Physics of Solids*, Vol. 32, pp. 461-490.
- 10 ROUSSELIER, G. (1987): "Ductile Fracture Models and Their Potential in Local Approach of Fracture", *Nuclear Engineering and Design*, Vol. 105, pp. 97-111.
- 11 ROSE, J. FERRANTE, J., AND SMITH, J. (1981): "Universal Binding Energy Curves for Metals in Bimetallic Interfaces", *Physical Review Letters*, Vol. 47, pp. 675-678.
- 12 NEEDLEMAN, A. (1992): "An Analysis of Decohesion along an Imperfect Interface", *Int. J. of Fracture*, Vol. 40, pp. 1377-1397.
- 13 YUAN, H., LIN, A., AND CORNEC, A. (1996): "Verification of a Cohesive Zone Model for Ductile Fracture", *ASME, J. Engineering and Technology*, Vol. 118, pp. 192-200.
- 14 SIEGMUND, T. AND BROCKS, W. (1999): "Prediction of the Work of Separation and Implications to Modeling", *Int. J. of Fracture*, Vol. 99, pp. 97-116.
- 15 SIEGMUND, T. AND BROCKS, W. (2000): "The Role of Cohesive Strength and Separation Energy for Modeling of Ductile Fracture", *Fatigue and Fracture Mechanics: 31st Vol.*, ASTM STP 1360, P.C. Paris and K.L. Jerina, Eds., pp. 139-151.
- 16 SIEGMUND, T., BROCKS, W., HEERENS, J., TEMPUS, G., AND ZINK, W. (1999): "Modeling of crack growth in thin sheet aluminium", *Recent Advances in Solids and Structures*, Proc. ASME Int. Mechanical Engineering Congress and Exposition (eds. Y.W. Kwon and H.C. Chung), Nashville, Nov. 1999, PVP 398, pp. 15-22.
- 17 SIEGMUND, T. AND BROCKS, W.: *Modeling Crack Growth in Thin Sheet Aluminum Alloys*, *Fatigue and Fracture Mechanics: 31st Vol.*, ASTM STP 1389, in press.

Address: PROF. WOLFGANG BROCKS, GKSS-Forschungszentrum, Max-Planck-Str., 21502 Geesthacht  
Germany



PETER GUMBSCH

## Atomistische Aspekte des Bruchs

*Die atomistische Modellierung von Rissen zielt darauf, die der Sprödigkeit von Werkstoffen zugrunde liegenden Prozesse verstehen zu lernen. Meist werden dabei einfache atomare Wechselwirkungsmodelle verwendet, in jüngster Zeit kommen aber auch quantenmechanische Methoden zum Einsatz. Atomistische Simulationen können eingesetzt werden, um mögliche Spaltebenen von Kristallen, deren Bruchzähigkeit sowie die Abhängigkeit von Rissfortschrittsrichtung oder Belastungsart zu ermitteln. Hierzu werden beispielhaft Ergebnisse zum Sprödbbruch von Wolfram und Silizium vorgestellt und mit experimentellen Untersuchungen verglichen.*

### 1. Einleitung

Der Bruch von Werkstoffen ist ein Phänomen, das über viele Längenskalen hinweg betrachtet werden muss. Die makroskopische Probengeometrie und die Belastungsart sind ebenso wichtig wie die Details der atomaren Bindungsverhältnisse an der Rissspitze. Die Bedeutung der atomaren Skala ist im Falle des perfekt spröden Bruchs offensichtlich, da sich ein Riss in einem spröden Material nur durch das Brechen atomarer Bindungen an der Rissspitze ausbreiten kann und daher atomar scharf sein muss. Halbspröde Materialien, wie die zentralen Übergangsmetalle und die Stähle, zeigen ein begrenztes Maß an Plastizität und eine ausgeprägte Lastratenabhängigkeit beim Bruch. Atomistische Prozesse sind hier entscheidend für den Wettbewerb zwischen dem Brechen der Bindungen und der Erzeugung von plastischer Verformung an der Rissspitze.

Da experimentelle Information auf der atomaren Längenskala kaum verfügbar ist, ist die atomistische Modellierung das einzige Werkzeug mit dem solche atomistischen Prozesse auf der passenden Zeit- und Längenskala untersucht werden können. Nach der Darlegung einiger generisch atomistischer Effekte soll in dieser Arbeit die Bedeutung der atomistischen Betrachtung am Beispiel der Spaltanisotropie bezüglich der Rissfortschrittsrichtung aufgezeigt werden.

Die Hauptschwierigkeit bei der atomistischen Modellierung von Bruchprozessen liegt im Aufbringen realistischer Randbedingungen. Dies kann beispielsweise durch die Kopplung von atomistischen und kontinuumsmechanischen Simulationen gelöst werden. Auf simulationstechnische Fragen kann hier nicht weiter eingegangen werden. Diesbezüglich wird auf die Literatur [1, 2] verwiesen. Anzumerken ist hier nur noch, dass die atomistische Modellierung im Vergleich zu den häufig eingesetzten Kontinuumsmethoden den Vorteil besitzt, dass sie Bruch als Folge einer äußeren Belastung selbstständig reproduziert, nachdem die Beschreibung der atomaren Wechselwirkung festgelegt wurde, während die Kontinuumsmethoden immer auf mehr oder minder detaillierte Versagenskriterien angewiesen sind.

### 2. Bedeutung der diskreten, atomistischen Natur der Werkstoffe

Die Behandlung spröder Bruchprozesse folgt traditionell der Betrachtung von Griffith [3], bei der für die Rissspitze ein thermodynamischer Gleichgewichtszustand gesucht wird. Das Griffith-Kriterium für die Stabilität eines Risses ergibt sich als Gleichgewicht zwischen der den Riss treibenden Kraft, der Energiefreisetzungsrate  $\mathcal{G}$ , sowie dem Risswiderstand des Materials  $\mathcal{R}$ , der Oberflächenenergie  $\gamma_s$  der zwei Bruchflächen [4]:

$$\mathcal{G} = \mathcal{R} = 2\gamma_s. \quad (1)$$

Die Energiefreisetzungsrate  $\mathcal{G}$  kann im Rahmen der Elastizitätstheorie, die den spröden Riss als Spannungssingularität der Stärke  $K$  beschreibt, berechnet werden [4]. Der Spannungsintensitätsfaktor  $K$  ergibt sich aus der Geometrie der Probe und der äußeren Last und führt im linear-elastischen Fall direkt zur Energiefreisetzungsrate

$$\mathcal{G} = \frac{K^2}{E'}, \quad (2)$$

wobei  $E'$  ein elastischer Modul ist. Das Stabilitätskriterium ist in einer solchen kontinuumsmechanischen Betrachtung gemäß Gl. (1) und (2) nur für die Griffith-Last  $K_G = \sqrt{2\gamma E'}$  erfüllt. Während eine solche linear-elastische

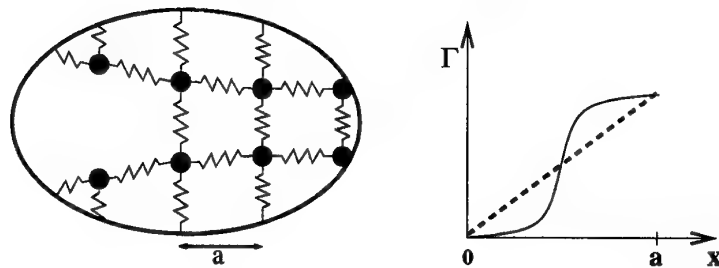


Abbildung 1: (Links) Schematische Darstellung einer Rissspitze im atomaren Gitter. (Rechts) Die Gesamtenergie der Oberfläche des Risses  $\Gamma$  steigt während des Rissfortschritts in der Kontinuumsbetrachtung mit der Steigung  $2\gamma_s$  linear an, während die atomistischen Verhältnisse eher durch einen nicht-linearen Anstieg beim Brechen der Bindung charakterisiert werden.

Beschreibung des Risses in einem spröden Material außerhalb der wenigen nicht-linearen Bindungen in unmittelbarer Umgebung der Rissspitze fraglos angemessen ist, muss die Annahme, der Materialwiderstand gegen Rissfortschritt wäre allein durch die Gibbs'sche Oberflächenenergie charakterisiert, näher beleuchtet werden. Atomistisch sollte der Materialwiderstand gegen Rissfortschritt durch die Kräfte charakterisiert sein, die notwendig sind, die atomaren Bindungen an der Rissspitze zu brechen. Die ersten atomistischen Betrachtungen zum Rissfortschritt [5] zeigten die Bedeutung der diskreten Natur der atomaren Bindungen darin, dass der Riss bis zu Lasten  $K^+ > K_G$  stabil bleibt und sich auch erst bei Lasten  $K^- < K_G$  wieder schließt, also in einem endlichen Lastbereich stabil ist. Das Vorhandensein eines solchen endlichen Stabilitätsbereichs bezeichnet man als *lattice trapping*. In einer simplen Kontinuumsanalogie könnte man das *lattice trapping* als nicht-linearen Anstieg der Oberflächenenergie beim Rissfortschritt betrachten (siehe Abb. 1).

Frühe atomistische Simulationsrechnungen zeigten [6], dass die Größe des *trapping*-Effekts stark von dem verwendeten Modell für die atomare Wechselwirkung abhängt. Werden die interatomaren Kräfte mit einem abrupt brechenden Federgesetz beschrieben, so kann  $K^+$  um einen Faktor zwei größer sein als  $K_G$ , während ein realistisches weiches und langreichweitiges Wechselwirkungspotential nur einen sehr geringen *trapping*-Effekt zeigt [6]. Weitergehende Betrachtungen [7] zeigen, dass die Größe des *trapping*-Effekts vorwiegend von der relativen Steifigkeit der Bindung an der Rissspitze und der diese Bindung belastenden Umgebung abhängt. In einem so generalisierten Bild ist das *trapping* auch nicht mehr auf ein atomares Gitter beschränkt, sondern lässt sich ebenso auf amorphe Festkörper oder Grenzflächen übertragen [8]. *Trapping*-Effekte können zwanglos verschiedene Beobachtungen, wie thermisch aktiviertes unterkritisches Risswachstum [9] oder die Erzeugung metastabiler Bruchflächen [2, 8], erklären. Ein besonders illustratives Beispiel ist aber die experimentell verschiedentlich beobachtete Spaltanisotropie bezüglich der Rissfortschrittsrichtung.

### 3. Richtungsanisotropie beim Spaltbruch

Im kontinuumsmechanischen Bild sollte ein Riss in einem elastisch isotropen Material auf einer Spaltebene richtungsunabhängig laufen können, sobald  $K_G$  erreicht ist, da die Oberflächenenergie in Gl. (1) keine Information über die Rissfortschrittsrichtung enthält und der Modul in Gl. (2) im isotropen Fall konstant ist. Elastische Anisotropie mag über Gl. (2) eine gewisse Richtungsanisotropie einbringen, diese ist aber üblicherweise schwach, da die Hauptbelastungsrichtung bei öffnender Belastung parallel zur Bruchflächennormale ist und diese natürlich nicht von der Fortschrittsrichtung abhängt. Im Gegensatz dazu wird im Experiment manchmal eine drastische Richtungsanisotropie beobachtet [10, 11, 12].

Der Spaltbruch von einkristallinem Wolfram, das elastisch nahezu isotrop ist, ist einer der am besten dokumentierten Fälle für eine solche Richtungsanisotropie. Atomistische Simulationsrechnungen [1] sowie Bruchexperimente bei  $T = 77\text{ K}$  [12] zeigen beide, dass Rissfortschritt auf beiden Spaltebenen, den  $\{100\}$ - und  $\{0\bar{1}1\}$ -Ebenen, leicht erreicht werden kann und defektarme Bruchflächen erzeugt, wenn die Rissfront parallel zu einer  $\langle 011 \rangle$ -Richtung orientiert ist. Im Gegensatz dazu werden etwa 40% höhere Belastungen benötigt um Rissfortschritt mit  $\langle 001 \rangle$ -Rissfronten zu erreichen. Die dabei erzeugten Bruchflächen sind außerdem rauher und zeigen Markierungen die andeuten, dass der Rissfortschritt lokal entlang der leichten Richtungen erfolgte [12]. Tabelle 1 fasst die experimentellen Ergebnisse und die berechneten Bruchzähigkeiten zusammen. Qualitativ ist die Übereinstimmung zwischen den 77 K-Experimenten und den Simulationsrechnungen ausgezeichnet, quantitativ sind die experimentell ermittelten Bruchzähigkeiten aber doch signifikant höher als die berechneten Werte. Dies könnte andeuten, dass die Rissspitzenplastizität selbst bei

Spaltsystem {Ebene}{Front}	Bruchexperiment Raumtemperatur	Bruchexperiment 77K	Atomistische Simulation	$K_G$ aus atomistischem Wechselwirkungsmodell
{100}{010}	$8.7 \pm 2.5$	$3.4 \pm 0.6$	2.05	1.61
{100}{011}	$6.2 \pm 1.7$	$2.4 \pm 0.4$	1.63	1.61
{110}{001}	$20.2 \pm 5.5$	$3.8 \pm 0.4$	2.17	1.51
{110}{110}	$12.9 \pm 2.1$	$2.8 \pm 0.2$	1.56	1.51

Tabelle 1: Bruchzähigkeit von Wolfram Einkristallen für {100}- und {110}-Spaltebenen für verschiedene Rissausbreitungsrichtungen/Rissfronten. Die experimentellen Daten stellen Mittelwert und Standardabweichung von mindestens 5 unabhängigen Messungen dar. Sie sind [2, 12] entnommen und in  $\text{MPa}\sqrt{\text{m}}$  angegeben.

77 K einen nicht vernachlässigbaren Einfluss hat. Der Vergleich von Raumtemperatur- und 77 K-Experiment zeigt zwar ein deutliches Absinken der plastischen Anteile an der Bruchzähigkeit mit sinkender Temperatur, dennoch kann deren Einfluss nicht ausgeschlossen werden.

Experimentell sind die Verhältnisse diesbezüglich im Silizium wesentlich klarer. Silizium kann als praktisch versetzungsfreier Einkristall hergestellt werden. In-situ Beobachtungen im Elektronenmikroskop [13] sowie in der Röntgentopographie [10, 11] zeigen, dass Risse im Silizium bei Raumtemperatur ohne jegliche begleitende plastische Verformung fortschreiten. Silizium bricht sowohl auf {111}- als auch auf {110}-Ebenen [10]. Beim Spaltbruch auf den {110}-Ebenen zeigt sich dabei eine besonders drastisch ausgeprägte Richtungsanisotropie. Während Risse sich entlang der  $\langle 110 \rangle$ -Richtung bei den niedrigsten gemessenen Bruchzähigkeiten ausbreiten und dabei nahezu perfekte ebene Bruchflächen erzeugen [10, 11], gelingt der Rissfortschritt in  $\langle 001 \rangle$ -Richtung nicht. Der Versuch, Risse in diese Richtung und damit senkrecht zur bevorzugten Ausbreitungsrichtung zu treiben, führt zu einem Abknicken des Risses und einer Ausbreitung auf {111}-Ebenen [11, 14]. Dies ist schematisch in Abb. 2 (links) dargestellt. Abb. 2 (rechts) zeigt die Seitenansicht eines makroskopisch entlang der  $\langle 001 \rangle$ -Richtung gebrochenen Siliziumwafers. Da für die {110}-Risse bei Ausbreitung entlang der leichten Richtungen eine niedrigere Bruchzähigkeit gemessen wird als für die {111}-Risse [10] und da die elastische Anisotropie im Silizium allenfalls für Unterschiede von 10% im effektiven Modul  $E'$  ausreicht [14], kann der beobachtete Unterschied konventionell nicht erklärt werden.

Obwohl sich ein solch ausgeprägter Effekt prinzipiell gut atomistisch modellieren lassen sollte, sind entsprechende Rechnungen mit einfachen interatomaren Wechselwirkungsmodellen nicht gelungen. Dies liegt daran, dass alle empirischen Wechselwirkungsmodelle für Silizium zur Stabilisierung der Kristallstruktur kurzreichweitig sein müssen und daher beim Brechen der Bindungen, wo die Wechselwirkungsmodelle an der Grenze ihrer Reichweite getestet werden, unangemessen steife Bindungen vorspielen. Diese Problematik ist bislang noch nicht befriedigend gelöst und der einzige Ausweg besteht darin, auf quantenmechanische Methoden zur Beschreibung der atomaren Wechselwirkung zurückzugreifen. Hierbei ist man jedoch bei der Anzahl zu behandelnder Atome sehr stark eingeschränkt, so dass der Riss immer signifikant mit den Randbedingungen wechselwirkt.

Selbstkonsistente quantenmechanische Berechnungen zum Brechen der Risspitzenbindungen im Silizium zeigen zwei qualitativ unterschiedliche Prozesse [14]: einerseits ein kontinuierliches Öffnen der Bindungen an {111}-Rissen oder

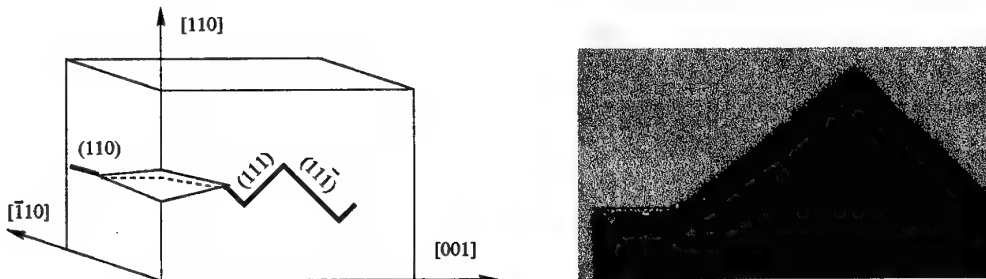


Abbildung 2: (Links) Schematische Darstellung der Richtungsanisotropie beim {110}-Spaltbruch von Silizium. Rissfortschritt entlang der  $\langle 110 \rangle$ -Richtung gelingt leicht und erzeugt eine nahezu defektfreie Bruchfläche, während senkrecht dazu der Riss von der {110}-Ebene auf {111}-Ebenen ausweicht. (Rechts) Die Seitenansicht einer makroskopisch in  $\langle 001 \rangle$ -Richtung gebrochenen Probe, die deutliche {111}-Facetten erkennen lässt. (Diese Abbildung wurde dankenswerterweise von T. Cramer zur Verfügung gestellt [15].)

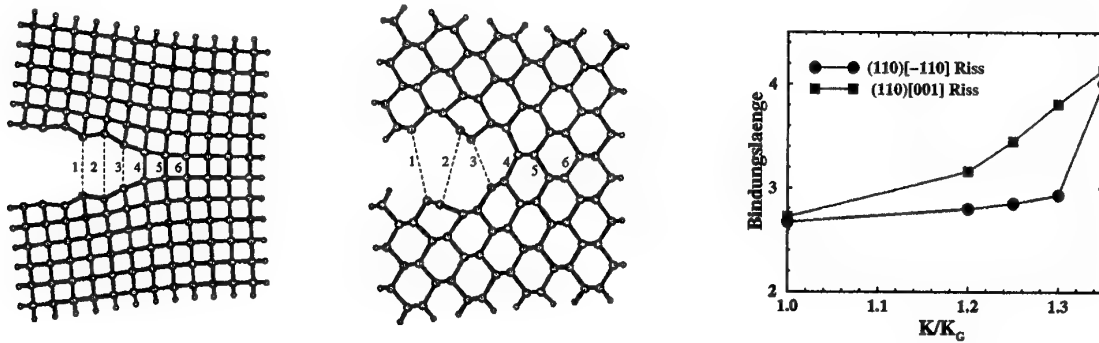


Abbildung 3: (Links) Mit selbstkonsistenten Pseudopotentialrechnungen relaxierte atomare Struktur des  $\{110\}\langle 001 \rangle$  Rissystems, welches experimentell leicht zu spalten ist, bei einer Belastung von  $1.2 K_G$ . (Mitte) Relaxierte Struktur des  $\{110\}\langle 110 \rangle$  Rissystems bei einer Belastung von  $1.2 K_G$ . (Rechts) Länge der Risspitzenbindung (# 4) für beide Rissysteme in Abhängigkeit von der aufgebrachten Belastung. Deutlich ist das abrupte Brechen der Bindung im  $\{110\}\langle 110 \rangle$  Rissystem bei einer Belastung von  $1.35 K_G$  zu erkennen. Die Bindungslänge ist in Ångström angegeben.

dem in der leichten Orientierung ausgerichteten  $\{110\}$ -Riss, andererseits ein abruptes Brechen für den  $\{110\}$ -Riss mit  $\langle 110 \rangle$ -Rissfront. Die atomare Struktur der Risspitze sowie die Abhängigkeit der Länge der Risspitzenbindung von der anliegenden Last sind in Abb. 3 dargestellt. Das kontinuierlich Öffnen der Bindung in der leichten Orientierung ist auf die sehr kleine Systemgröße zurückzuführen, weist aber sicher auf geringe *trapping*-Effekte hin. Im Gegensatz dazu ist das abrupte Brechen unabhängig von der Systemgröße. Es geht mit einem signifikanten *trapping* einher und hat seine Ursache in einer Aufteilung der Last zwischen den zwei unmittelbar an der Risspitze liegenden Bindungen (siehe Abb. 3) [14]. Aufgrund des großen *trapping*-Effekts in der schwierigen Fortschrittsrichtung ist es für den Riss möglich auf die mit geringem *trapping* zugänglichen  $\{111\}$ -Ebenen auszuweichen. Die atomistischen Rechnungen können damit die experimentellen Beobachtungen zumindest qualitativ erklären.

Zusammenfassend kann damit festgehalten werden, dass atomistische Betrachtungen zur Analyse spröder Bruchprozesse sehr hilfreich sind und mit ihnen insbesondere auch die Bevorzugung einzelner Rissfortschrittsrichtungen gut erklärt werden kann.

#### 4. Literatur

- 1 KOHLHOFF, S., GUMBSCH, P., FISCHMEISTER, H.F.: Crack propagation in b.c.c. crystals studied with a combined finite-element and atomistic model. *Phil. Mag. A* **64** (1991) 851.
- 2 GUMBSCH, P.: Brittle Fracture Processes Modelled on the Atomic Scale. *Z. Metallkd.* **87** (1996) 341.
- 3 GRIFFITH, A. A.: The Phenomena of Rupture and Flow in Solids. *Philos. Trans. R. Soc.* **221A** (1921) 163.
- 4 THOMSON, R.: Physics of Fracture. In *Solid State Physics*, edited by H. Ehrenreich and D. Turnbull (Academic Press, New York, 1986) 1.
- 5 THOMSON, R., HSIEH, C., RANA, V.: Lattice Trapping of Fracture Cracks. *J. Appl. Phys.* **42** (1971) 3154.
- 6 SINCLAIR, J. E.: The influence of the interatomic force law and of kinks on the propagation of brittle cracks. *Phil. Mag.* **31** (1975) p. 647.
- 7 SCHOECK, G., PICHL, W.: Bond Trapping of Cracks. *phys. stat. sol. (a)* **118** (1990) 109.
- 8 GUMBSCH, P., CANNON, R. M.: Atomistic Aspects of Brittle Fracture. *MRS Bulletin* 25/5 (2000) 15-20.
- 9 COOK, R.F.: Crack propagation thresholds: A measure of surface energy. *J. Mater. Res.* **1** (1986) p. 852.
- 10 MICHOT, G.: Fundamentals of Silicon Fracture. *Crystal Prop. and Prep.* **17** (1988) 55.
- 11 GEORGE, A., MICHOT, G.: Dislocation loops at crack tips: nucleation and growth – an experimental study in silicon. *Mat. Sci. Eng. A* **164** (1993) 118.
- 12 RIEDLE, J., GUMBSCH, P., FISCHMEISTER, H.F.: Cleavage Anisotropy in Tungsten Single Crystals. *Phys. Rev. Letters* **76** (1996) 3594.
- 13 CLARKE, D.: Fracture of Silicon and Other Semiconductors. In *Semiconductors and Semimetals*, Vol. 37 edited by Katherine T. Faber (Academic Press, New York, 1992) 79.
- 14 PÉREZ, R., GUMBSCH, P.: Directional Anisotropy in the Cleavage Fracture of Silicon. *Phys. Rev. Letters* **84** (2000) 5347.
- 15 CRAMER, T.: Experimentelle Untersuchung zur dynamischen Rissausbreitung in spröden Materialien. Dissertation, Universität Stuttgart, 1999.

*Anschrift:* Dr. Peter Gumbsch, Max-Planck-Institut für Metallforschung, Seestr. 92, 70174 Stuttgart, Germany.

Agarwal, R. K.

## Acoustic Radiation Due to Gust-Airfoil and Blade-Vortex Interactions

*An accurate and efficient method for computing acoustic radiation due to gust-airfoil and blade-vortex interactions is developed. In these types of problems, sound is generated as a result of interaction between the unsteadiness in the flow and the body. The acoustic governing equations are derived by linearizing the compressible unsteady Euler equations about the steady mean flow. From these equations, the frequency domain acoustic equations are obtained assuming a single frequency disturbance. The equations are solved by employing a multi-stage Runge-Kutta finite-volume time-stepping scheme with a fourth-order compact spatial discretization. In the farfield, both the Giles' nonreflecting boundary condition and the Perfectly Matched Layer (PML) absorbing boundary conditions are employed. This report describes the technical approach and shows the results calculated for the interactions.*

### 1. Introduction

An issue of importance to aircraft designers and manufacturers is the accurate prediction of aerodynamically generated noise. Until the last decade, most of the analytical and computational work in aeroacoustics has been based on the wave equation analogy type of approaches pioneered by Lighthill [1]. But with the maturation of Computational Fluid Dynamics (CFD) technology over the last two decades, it is now feasible to develop acoustic codes based on the solution of unsteady compressible Euler/Navier-Stokes equations. However, since the nature and characteristics of aeroacoustics problems are different from those encountered in aerodynamics, CFD based Computational Aeroacoustics (CAA) technology needs to resolve these characteristics in an accurate and efficient manner. The author and his colleagues have developed an accurate and efficient numerical method for solving a wide variety of problems in acoustic propagation, radiation and scattering [2]. This paper deals with the application of the acoustic code described in [2] to the problems of acoustic radiation due to gust-airfoil and blade-vortex interactions.

### 2. Technical Approach

The time-domain acoustic equations are derived from the unsteady compressible Euler equations by linearizing about a steady mean flow. From these equations, the frequency-domain acoustic equations are obtained by assuming a single frequency disturbance. A pseudo-time variable is introduced into the frequency-domain equations so that the same computational techniques that are used to integrate the time-domain equations in time can be employed to time-march the frequency-domain equations to harmonic state. This approach provides a unified framework for the solution of both time- and frequency-domain acoustic equations. The equations are solved by employing a multi-stage Runge-Kutta finite-volume time-stepping algorithm with a fourth-order compact spatial discretization and a six-order compact filter (dissipation). A rigid wall type of boundary condition is applied in all calculations such that the total normal velocity on the airfoil is zero. In the farfield, both the Perfectly Matched Layer (PML) absorbing boundary condition [3] and Giles non-reflecting boundary condition [4] are employed.

### 3. Acoustic Radiation due to Gust-Airfoil Interaction

We consider an airfoil in a steady compressible flow when it encounters an unsteady gust. For simplicity, the scope of investigation is limited to an incoming sinusoidal gust such that the transverse velocity perturbation can be modeled as  $v' = e^{i(kx - \omega t)}$ , where  $k = \omega / u_\infty$  and  $u_\infty$  is the free stream

velocity. The gust/flat-plate interaction benchmark problem from the first ICASE/LaRC Computational Aeroacoustics Workshop [5] was chosen for computations. It is a relatively high-frequency non-compact case with reduced frequency  $k = 11.8$  and compactness ratio  $K = 7.85$ . The sensitivity of the solution to the numerical parameters such as the order of the difference method, the grid size, the farfield boundary location and the level of artificial dissipation were examined. PML absorbing boundary condition and the sixth-order Padé scheme were used unless noted otherwise. The features of the acoustic pressure field for this non-compact gust/flatplate interaction are illustrated in Figures 1 and 2 at Mach numbers of 0.25 and 0.75 respectively. The Doppler shift in the propagating waves is apparent in the instantaneous pressure contours. Results for the farfield pressure obtained with the fourth-, sixth- and eighth-order Padé schemes at Mach number of 0.5 are shown in Figure 3. The differences in the solutions are insignificant indicating that the fourth-order method has sufficient resolution for this case. The results compare well with the solutions published in Reference 5. Computed plate surface pressure distributions are shown in Figure 4. As the compactness ratio is increased, there is an increase in the number of lobes in the sound pattern as a result of phase differences and variable loading along the plate surface. We have also computed the problem of gust-airfoil interaction for a compact case with  $K = 0.5$ , a NACA0001 airfoil with freestream Mach number of 0.5,  $\omega = 0.8875$  (40 Hz) and a gust of unit amplitude. Figure 5 compares the amplitude of the computed acoustic pressure with asymptotic expansion of Amiet [6] along the lower surface of the flat plate. The agreement is excellent.

#### 4. Acoustic Radiation due to Blade-Vortex Interaction

A significant source of helicopter noise is the acoustic radiation due to blade-vortex-interaction (BVI). When a vortex is shed from the rotating blade tip and is convected downstream, it interacts with the next blade which results in the movement of the leading edge stagnation point on the blade resulting in change in the lift. The rate of change of lift induces pressure fluctuations in the proximity of the blade, and results in a series of expansion/compression waves propagating upstream. Research has shown that when the blade is parallel to the vortex axis, the noise due to BVI is very high and the interaction can be modeled as 2-D. In this investigation, a finite core Rankine vortex embedded in the flow field was used to simulate a shedding vortex. The vortex core is a rotational region and contains all the vorticity associated with the vortex. In order to study the unsteady inviscid interaction of a Rankine vortex with an airfoil, the steady-state flow past the airfoil was first computed using a CFD code. A Rankine vortex was then introduced into the flowfield at a point upstream of the airfoil. Time-accurate computations were started from this initial condition. The vortex was allowed to freely interact with the airfoil so that subsequent convection required no explicit tracking. Using the magnitudes of the perturbation qualities (due to the interaction of vortex with airfoil) determined from CFD, the acoustic signature was computed with the frequency-domain acoustic solver with the frequency determined either from BVI experiments, or by performing the Fast Fourier Transform (FFT) analysis of the unsteady pressure time history obtained from the CFD solver. Parametric studies of vortex location, vortex strength, and flow field conditions on unsteady pressure field about various airfoil sections were conducted. The case of a NACA0012 airfoil interacting with a vortex of strength 0.2 at convective Mach number ranging from 0.3 to 0.8 was considered. Aerodynamic characteristics due to the BVI at Mach 0.8 are shown in Figure 6 which compare well with the solutions published in Reference 7. At a frequency of 200Hz, the variation of calculated acoustic intensity with distance along a 45-degree ray below the airfoil leading edge at various subsonic freestream Mach number is shown in Figure 7. In a similar fashion, the acoustic intensity for two different values of vortex strength at Mach 0.3 is plotted in Figure 8. These calculations are similar to those reported by Baeder [8] et al. The computations show that higher values of acoustic intensity are obtained for higher Mach number, stronger vortex strength, smaller vortex core size, or smaller distance between the vortex and the airfoil.

#### 5. Conclusions

A fourth- order compact time-domain/frequency-domain computational acoustics code is applied to compute the acoustic radiation due to gust-airfoil and blade-vortex interactions. Excellent solutions are obtained which compare well with the analytical solutions or the computations of other investigators. Extensive calculations show that a fourth-order Padé scheme with PML absorbing



boundary conditions in the farfield provides the most accurate and efficient method for acoustic computations using linearized Euler equations.

## 6. References

- 1 Lighthill, M.J.: On Sound Generated Aerodynamically – I. General Theory; Proc. Roy. Soc., **211A** (1952), 564-587.
- 2 AGARWAL, R.K. and Huh, K.S., Scattering of Sound by Rigid Bodies in Arbitrary Flows; CFD Review (1995), M. Hafez and K. Oshima Editors, John Wiley (1995), 797-820.
- 3 GILES, M., Non-Reflecting Boundary Conditions for Unsteady Airfoil Calculations; CFDL-TR-90-1; MIT Department of Aeronautics and Astronautics, 1990.
- 4 HU, F.Q., On Absorbing Boundary Conditions for Linearized Euler Equations by a Perfectly Matched Layer; Journal of Computational Physics, **129** (1996), 201-219.
- 5 HARDIN, J.C., RISTORCELLI, J.R., and TAM, C.K.W. (eds.), ICASE/LaRC Workshop on Benchmark Problems in Computational Aeroacoustics (CAA); NASA CP-3300, 1995.
- 6 AMIET, R.K., Compressibility Effects in Unsteady Thin-Airfoil Theory; AIAA Journal, **12**, No. 2 (1974), 252-255.
- 7 DAMODARAN M. and CAUGHEY, D. A., Finite-Volume Calculation of Inviscid Transonic Airfoil-Vortex Interaction; AIAA Journal, **26** No. 11 (1988) 1346-1353.
- 8 BAEDER, J.D., MCCRASKEY, W.J. and SVINIVASAN, G., Acoustic Propagation Using Computational Fluid Dynamics; Proc. of AMS Annual Forum, **1** (1986), 551-562.

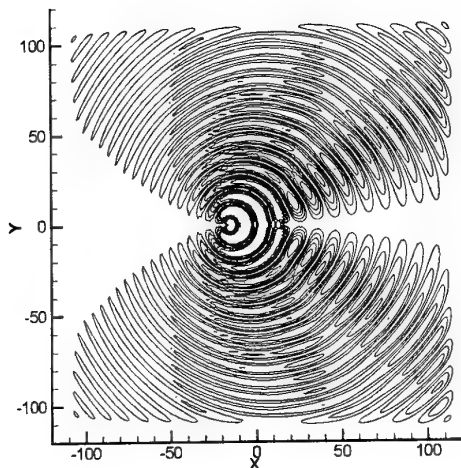


Figure 1. Instantaneous Pressure Field for Noncompact Gust at  $M = 0.25$

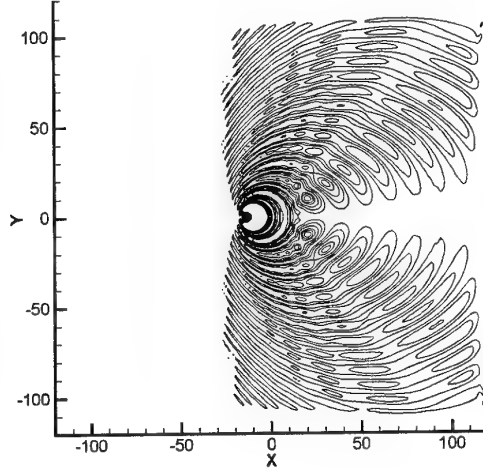


Figure 2. Instantaneous Pressure Field for Noncompact Gust at  $M = 0.75$

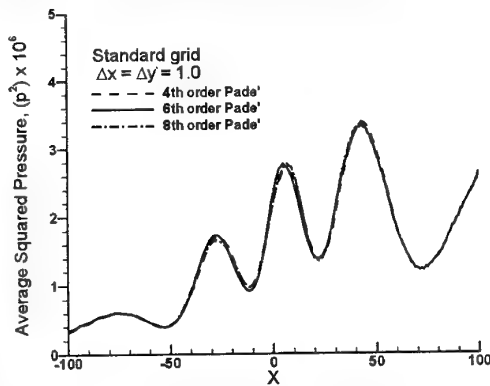


Figure 3. Comparison of Pressure Profiles at  $y = 95$  for the High-Order Compact Methods

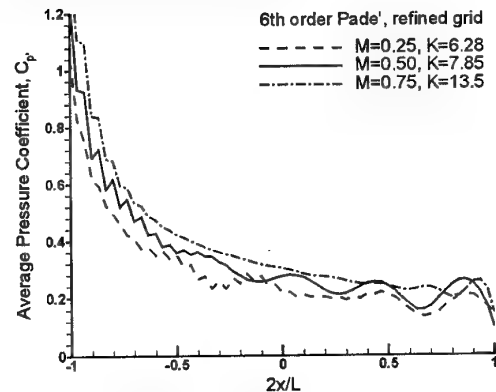


Figure 4. Surface Pressure Distribution for  $\omega = 2\pi/16$

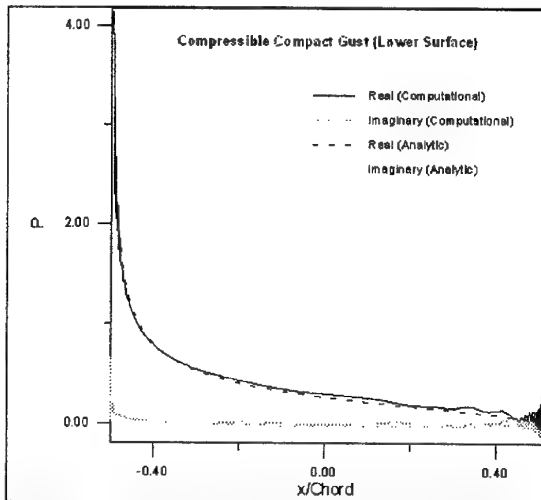


Figure 5. Airfoil Surface Acoustic Pressure

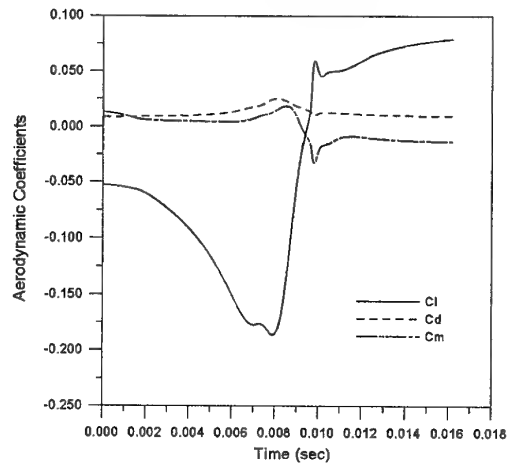


Figure 6. BVI CFD Results at Mach 0.8

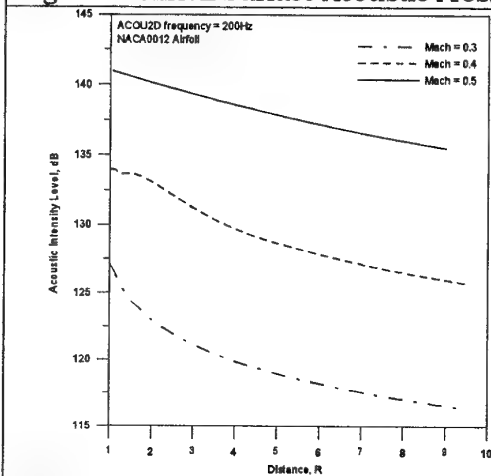


Figure 7. BVI Acoustic Intensity for Various Freestream Mach Numbers

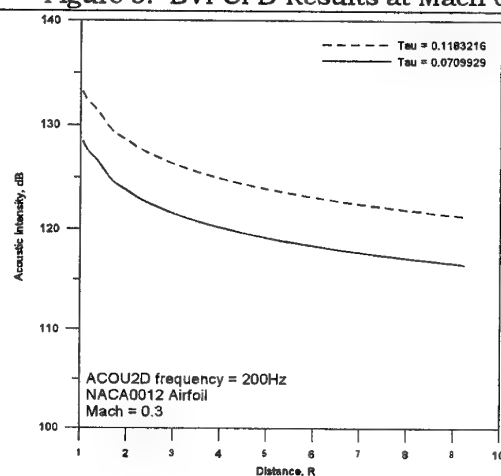


Figure 8. BVI Acoustic Intensity for Various Vortex Strengths

Address: Dr. Ramesh K. Agarwal, National Institute for Aviation Research, Wichita State University, Wichita, Kansas 67260-0093, USA



BLACKMORE, D; KNIO, O.

## Hamiltonian Structure for Vortex Filament Flows

*A new Hamiltonian formulation of the passive particle motion induced by a smooth vortex filament in an ideal fluid contained in a region of 3-space is derived. The point of departure in the derivation is a desingularized version of the Biot-Savart formula for the induced velocity field. Then a foliation of a neighborhood of the filament (that moves with the fluid flow) is constructed that is comprised of smooth two-dimensional leaves that are invariant with respect to the induced velocity field at each time. Natural symplectic coordinates are introduced on the moving leaves of the associated foliation such that the equations of motion on the leaves assume a simple (possibly time-dependent) Hamiltonian form. With this Hamiltonian structure one can, by simply following the evolution of the leaves of the foliation, easily determine the motion of the passive fluid particles near the filament. Any irregular or singular behavior in the motion can essentially be associated to geometrical features of the moving foliation in the large. The Hamiltonian structure is illustrated with three examples: a rectilinear filament; a circular vortex ring; and a helical filament.*

### 1. Introduction

Let  $C_t$  be a smooth vortex filament having constant vortex strength  $\Gamma \neq 0$ , moving in an ideal fluid in a region of  $\mathbb{R}^3$  for time  $t$  ( $t > 0$ ). We shall show how to introduce natural (moving) coordinates  $(q, p)$  transverse to  $C_t$  such that the fluid motion induced by the filament can be described in the Hamiltonian form

$$\dot{q} = \{q, H\} = \frac{\partial H}{\partial p}(q, p, t), \quad \dot{p} = \{p, H\} = -\frac{\partial H}{\partial q}(q, p, t), \quad (1)$$

in a neighborhood  $\mathfrak{N}_t$  of  $C_t$  for all  $t \geq 0$ , where the  $\dot{\cdot}$  denotes  $d/dt$  and  $\{\cdot, \cdot\}$  is the usual canonical Poisson bracket.

The idea is to first construct a smooth family of foliations  $\mathcal{F} := \{\mathfrak{F}_t : t \geq 0\}$  such that  $\mathfrak{F}_t$  is a foliation of  $\mathfrak{N}_t$  comprised of leaves that are smooth surfaces transverse to  $C_t$  and each leaf  $\mathfrak{L}$  of  $\mathfrak{F}_t$  is invariant with respect to the induced velocity field for each  $t \geq 0$ . We then introduce natural, canonical (moving) coordinates on the leaves of  $\mathcal{F}$  that produce the desired Hamiltonian structure. Our approach is somewhat like, but more intuitive and geometric than, the Clebsch transformation (see [6, 7, 11]) - which actually is not strictly applicable to the flows under consideration.

We develop the subject matter of this paper in the following manner: In Section 2 we delineate our basic assumptions concerning the nature of the vortex filament and the fluid flow. Next, in Section 3, we construct a moving foliation on whose leaves the relative fluid motion may be expressed in Hamiltonian form. Then in Section 4 we show how the relative Hamiltonian structure of Section 3 leads directly to the Hamiltonian equation (1). In Sections 5, 6 and 7, respectively, we apply our procedure for the construction of a Hamiltonian structure to a rectilinear vortex filament, a circular vortex ring and a helical vortex filament. Finally, we discuss our results in Section 8.

### 2. Preliminaries

We assume that the vortex filament has the form

$$C_t := \{\mathbf{R}(\sigma, t) = (\xi(\sigma, t), \eta(\sigma, t), \zeta(\sigma, t)) : (\sigma, t) \in I \times \mathbb{R}_+\}, \quad (2)$$

with all of the coordinate functions smooth. The motion of the filament as well as all other fluid particles is assumed to be governed by the following desingularized version of the Biot-Savart law:

$$\mathbf{v} = \mathbf{v}(\mathbf{r}, t) = \frac{\Gamma}{4\pi} \int_I \nabla (\chi(\|\mathbf{r} - \mathbf{R}(\sigma, t)\|)) \times \frac{\partial \mathbf{R}}{\partial \sigma} d\sigma, \quad (3)$$

where  $\nabla$  is the standard gradient operator (with respect to the space variables only),  $I$  is the interval on the real line on which the parameter  $\sigma$  of the filament is defined,  $\mathbf{r} = (x, y, z)$  and

$$\chi(\|\mathbf{r} - \mathbf{R}\|) := \psi(\|\mathbf{r} - \mathbf{R}\|) \|\mathbf{r} - \mathbf{R}\|^{-1} + (1 - \psi(\|\mathbf{r} - \mathbf{R}\|)) \log \|\mathbf{r} - \mathbf{R}\|^{-1}. \quad (4)$$

Here  $\psi : \mathbb{R} \rightarrow \mathbb{R}$  is a smooth function such that  $\psi(s) = 0$  for  $s \leq 0$ ,  $\psi(s) = 1$  for  $s \geq \delta > 0$  and  $d\psi/ds \geq 0$  for all  $s$ , where  $0 < \delta \ll 1$  represents the core radius of the vortex filament. This desingularization has the same essential features as those derived from fundamental principles for matching Navier-Stokes flows (e.g., see [2, 8, 9]). Note that if the filament has infinite length, integrability considerations may require an adjustment of the integrand in (3).

In order to evaluate the velocity field defined by (3), it is necessary to compute the parametrization  $\mathbf{R}(\sigma, t)$  of the moving filament which is determined by solving the partial differential equation

$$\frac{\partial \mathbf{R}}{\partial t}(\nu, t) = \frac{\Gamma}{4\pi} \int_I \nabla (\chi(\|\mathbf{R}(\nu, t) - \mathbf{R}(\sigma, t)\|)) \times \frac{\partial \mathbf{R}}{\partial \sigma}(\sigma, t) d\sigma, \quad (5)$$

where  $\mathbf{R}(\nu, 0)$  is any convenient parametrization of the initial filament  $C_0$ . We remark that (5) is an infinite-dimensional Hamiltonian dynamical system (in the same sense as the Euler equations) - one which we conjecture to be completely integrable (cf. [3, 4]). It is convenient to denote the "(semi-) flow" generated by  $\mathbf{v}$  as  $\varphi_t$  for  $t \geq 0$ . Then, for example, we can write  $C_t = \varphi_t(C_0)$ .

We shall construct a (moving) coordinate system  $(q(\mathbf{r}, t), p(\mathbf{r}, t), \omega(\mathbf{r}, t))$  adapted to  $C_t$  and its corresponding foliation  $\mathcal{F}_t$  for  $t \geq 0$  such that the leaves of the foliation are of the form  $\mathcal{L}_t^c = \{\mathbf{r} \in \mathbb{R}^3 : \omega(\mathbf{r}, t) = c\}$ , where  $c$  is a real constant,  $q$  and  $p$  are coordinates on the leaves and the motion on each leaf for  $t$  fixed is governed by

$$q' = \frac{\partial H^t}{\partial p}(q, p), \quad p' = -\frac{\partial H^t}{\partial q}(q, p), \quad (6)$$

where the prime denotes  $d/d\tau$ . Each leaf  $\mathcal{L}_t^c$  is  $\mathbf{v}(\cdot, t)$ -invariant and  $\tau$  parametrizes the trajectories of the motion on the leaves that would be generated by the velocity field if it were fixed in time. Following the coordinate system as it evolves under the action of  $\varphi_t$ , we shall find that the motion of passive fluid particles can be described in terms of the Hamiltonian system (1). For an observer at the origin of a leaf  $\mathcal{L}_t^c$ , which coincides with the point of intersection of the filament with the leaf, the motion appears two-dimensional (= one-degree-of-freedom) and Hamiltonian.

### 3. Derivation of Relative Hamiltonian Structure

For each  $t \geq 0$  we define  $\rho(\mathbf{r}, t)$  to be the distance from  $\mathbf{r}$  to  $C_t$  and  $\mathfrak{N}_t(\epsilon) := \rho^{-1}([0, \epsilon], t)$  to be the tubular  $\epsilon$ -neighborhood of  $C_t$ , where  $\epsilon > 0$ . Now we fix  $t \geq 0$ . Observe that it follows from (3) that  $\operatorname{div} \mathbf{v} = 0$  and that  $\mathbf{v} = \operatorname{curl} \mathbf{w}$  for all  $(\mathbf{r}, t) \in \mathbb{R}^3 \times \mathbb{R}_+$ , where

$$\mathbf{w} = \mathbf{w}(\mathbf{r}, t) := \frac{\Gamma}{4\pi} \int_I \chi(\|\mathbf{r} - \mathbf{R}(\sigma, t)\|) \frac{\partial \mathbf{R}}{\partial \sigma} d\sigma. \quad (7)$$

Consequently, by employing some standard results from vector analysis we infer that there exist smooth, real-valued functions  $\alpha$  and  $\omega$ , with  $\omega$  multi-valued if  $C_0$  is closed, such that

$$\mathbf{w} = \alpha \nabla \omega \implies \mathbf{v} = \nabla \alpha \times \nabla \omega \quad (8)$$

on  $\mathfrak{N}_t(\epsilon) \times \mathbb{R}_+$  for  $\epsilon$  sufficiently small. Observe that  $\mathbf{w}$  is tangent to  $C_t$  when  $\mathbf{r} \in C_t$ , the level surfaces of  $\omega$  for  $t \geq 0$  yield the leaves of the desired family of foliations  $\mathcal{F}$  and  $\alpha$  can serve as a coordinate in the leaves of  $\mathcal{F} := \{\mathcal{F}_t\}$ .

Rather than using  $\alpha$  as a coordinate in the leaves, we shall construct smooth functions  $\beta$  and  $\lambda$  such that  $(\beta, \lambda, \omega)$  are orthogonal coordinates for  $\mathfrak{N}_t(\epsilon)$  for all  $t \geq 0$  if  $\epsilon$  is sufficiently small. We draw curves along a fixed leaf  $\mathcal{L}_t^c$  that are parallel to the principal normal of  $C_t$  at the origin of the leaf. Then we flow these curves along  $\mathbf{w}$ . Thus we obtain a foliation of  $\mathfrak{N}_t(\epsilon)$  by surfaces that can be represented in the form  $\beta(\mathbf{r}, t) = \text{const.}$  if  $\epsilon$  is suitably small. Note that  $\nabla \beta \cdot \nabla \omega \equiv 0$  owing to definition of  $\beta$ . Analogously, by drawing a family of curves on  $\mathcal{L}_t^c$  that are orthogonal to the  $\beta$ -curves and flowing them along  $C_t$ , we obtain another foliation of  $\mathfrak{N}_t(\epsilon)$  by surfaces of the form  $\lambda(\mathbf{r}, t) = \text{const.}$  such that  $\{\nabla \beta, \nabla \lambda, \nabla \omega\}$  constitute an orthogonal basis on  $\mathfrak{N}_t(\epsilon)$  for all  $t \geq 0$  when  $\epsilon$  is sufficiently small. In terms of these orthogonal coordinates, we readily compute that

$$\mathbf{r}' = \frac{\partial \mathbf{r}}{\partial \beta} \beta' + \frac{\partial \mathbf{r}}{\partial \lambda} \lambda' + \frac{\partial \mathbf{r}}{\partial \omega} \omega' \implies \omega' \equiv 0 \implies \operatorname{div} \mathbf{v} = \kappa^{-1} \left[ \frac{\partial}{\partial \beta} (\kappa \beta') + \frac{\partial}{\partial \lambda} (\kappa \lambda') \right] = 0, \quad (9)$$

where  $\kappa := \|\partial \mathbf{r} / \partial \beta\| \|\partial \mathbf{r} / \partial \lambda\| \|\partial \mathbf{r} / \partial \omega\|$ , owing to the facts that  $\mathbf{r}' = \mathbf{v}$ ,  $\mathbf{v} \cdot \nabla \omega = 0$  and  $\operatorname{div} \mathbf{v} = 0$ . We infer from (9) that for each  $t \geq 0$  there exist a smooth (stream) function  $H^t$ , depending only on  $\beta$  and  $\lambda$ , such that  $\kappa \beta' = \partial H^t / \partial \lambda$

and  $\kappa\lambda' = -\partial H^t/\partial\beta$ . To obtain the desired Hamiltonian form on the (moving) leaves, we introduce the new variables

$$q := \int_0^\beta \kappa(s, \lambda) ds, \quad p := \lambda; \quad (10)$$

this leads directly to (6).

#### 4. Derivation of Absolute Hamiltonian Structure

Now that we have derived the relative Hamiltonian structure (6) of the flows, it is a simple matter to obtain the absolute Hamiltonian equations (1). We continue to use the orthogonal coordinates  $(\beta, \lambda, \omega)$  constructed in the preceding section. Then the reasoning used above can be employed to obtain the following absolute analog of (9):

$$\mathbf{v} = \dot{\mathbf{r}} = \frac{\partial \mathbf{r}}{\partial \beta} \dot{\beta} + \frac{\partial \mathbf{r}}{\partial \lambda} \dot{\lambda} + \frac{\partial \mathbf{r}}{\partial \omega} \dot{\omega} \implies \dot{\omega} \equiv 0 \implies \operatorname{div} \mathbf{v} = \kappa^{-1} \left[ \frac{\partial}{\partial \beta} (\kappa \dot{\beta}) + \frac{\partial}{\partial \lambda} (\kappa \dot{\lambda}) \right] = 0. \quad (11)$$

Hence there exists a smooth function  $H(\beta, \lambda, t)$ , where we may suppress the dependence on  $\omega$  since it is constant along the flow  $\varphi_t$  owing to (11). As  $\operatorname{div} \mathbf{v} = 0$  and  $\kappa = (\nabla \beta \cdot \nabla \lambda \times \nabla \omega)^{-1}$ ,  $\kappa$  is independent of  $t$  along  $\varphi_t$ . Consequently, we may proceed from this point just as in Section 3 to define the new variables (10) in terms of which the equations of motion take the form (1). Note that  $H$  in (1) and  $H^t$  in (6) are related as follows:  $H^t(q, p) = H(q, p, t)$ .

#### 5. Rectilinear Vortex Filament

For the case of a straight line vortex filament we may assume that  $C_0 = \{(0, 0, z) : z \in \mathbb{R}\}$ . It is easy to verify in this case that  $\mathbf{R}(z, t) = \mathbf{R}(z, 0)$  and  $\mathfrak{F}_t = \mathfrak{F}_0 = \{\text{the foliation of } \mathbb{R}^3 \text{ by planes parallel to the } x, y\text{-plane}\}$  for all  $t \geq 0$ . We find that  $\beta = x$ ,  $\lambda = y$ , and  $\omega = z$  comprise a suitable global coordinate system for the Hamiltonian structure. As the structure is independent of  $t$ , the relative and absolute motion coincide and we compute that the desired equations on the leaves are

$$\dot{q} = \frac{\partial}{\partial p} \left( \frac{\Gamma}{4\pi} \int_{-\infty}^{\infty} \chi(\sqrt{q^2 + p^2 + s^2}) ds \right), \quad \dot{p} = -\frac{\partial}{\partial q} \left( \frac{\Gamma}{4\pi} \int_{-\infty}^{\infty} \chi(\sqrt{q^2 + p^2 + s^2}) ds \right), \quad (12)$$

where  $q := x$  and  $p := y$ . Note that in this case  $\chi$  must be adjusted, for example by the factor  $\exp(-\sqrt{q^2 + p^2 + s^2})$ , in order to insure the existence and smoothness of the Hamiltonian function in (12).

#### 6. Circular Vortex Ring

The initial circular vortex ring may be assumed to be in the form  $C_0 = \{\mathbf{R}(\phi, 0) = (a \cos \phi, a \sin \phi, 0) : 0 \leq \phi \leq 2\pi\}$ , where  $a > 0$  is the initial radius of the ring. It is easy to show that  $\mathfrak{F}_t = \mathfrak{F}_0 = \{\text{the singular foliation (with singularity the } z\text{-axis) of } \mathbb{R}^3 \text{ by half-planes } \phi = \text{const.}\}$  for all  $t \geq 0$ . In this case we readily compute that  $C_t = \{\mathbf{R}(\phi, t) = (a \cos \phi, a \sin \phi, \gamma t) : 0 \leq \phi \leq 2\pi\}$ , where  $\gamma$  is the nonzero constant defined by

$$\gamma := \frac{\Gamma}{4\pi} \left\| \int_0^{2\pi} \nabla \left( \chi(a\sqrt{2(1-\cos\phi)}) \right) \times \frac{\partial \mathbf{R}}{\partial \phi}(\phi, 0) d\phi \right\|, \quad (13)$$

Note that the improper integral in (13) is convergent. A coordinate system for the structure is  $\beta = \rho := (x^2 + y^2)^{1/2}$ ,  $\lambda = z$ ,  $\omega = \phi$ . Whence, the Hamiltonian equations in terms of  $q := \rho^2$  and  $p := z$  are readily found to be

$$\dot{q} = \frac{\partial H}{\partial p}, \quad \dot{p} = -\frac{\partial H}{\partial q}; \quad H = -\frac{a\Gamma\sqrt{q}}{\pi} \int_0^\pi \chi \left( \sqrt{(\sqrt{q}-a)^2 + (p-\gamma t)^2 + 4a\sqrt{q}\sin^2 \frac{\phi}{2}} \right) \cos \phi d\phi, \quad (14)$$

#### 7. Helical Vortex Filament

A circular helical filament can be parametrized as  $C_0 = \{\mathbf{R}(\sigma, 0) = (a \cos \sigma, a \sin \sigma, b\sigma) : a, b > 0, \sigma \in \mathbb{R}\}$ . We find after a straightforward calculation that  $C_t = \{\mathbf{R}(\sigma, t) = (a \cos \sigma, a \sin \sigma, b\sigma + \mu t)\}$ , where  $\mu$  is the nonzero constant

defined as

$$\mu := \frac{a\Gamma}{4\pi} \int_{-\infty}^{\infty} \left[ \cos \sigma \frac{\partial \chi}{\partial x} \left( \sqrt{2a^2(1 - \cos \sigma)^2 + b^2 \sigma^2} \right) + \sin \sigma \frac{\partial \chi}{\partial y} \left( \sqrt{2a^2(1 - \cos \sigma)^2 + b^2 \sigma^2} \right) \right] d\sigma. \quad (15)$$

Note that as in Section 5,  $\chi$  requires an adjustment to insure smoothness of the integrals in this section.  $\mathfrak{F}_t$  changes with time: its leaves take the form of the helicoids  $\mathfrak{L}_t^c = \{\omega(\mathbf{r}, t) := (z - \mu t) - a \tan^{-1}(y/x) = c\}$ . It is not difficult to show that  $(\beta, \lambda, \omega)$  is a suitable orthogonal coordinate system, where  $\beta := \rho$  and  $\lambda := (z - \mu t) - b \tan^{-1}(y/x)$ . Defining  $q := \lambda^2$  and  $p := \beta$ , we find after a complicated but routine calculation that the desired Hamiltonian equations of motion on the leaves are (cf. [8]):

$$\dot{q} = \frac{\partial H}{\partial p}, \quad \dot{p} = -\frac{\partial H}{\partial q}; \quad H := -\frac{a\Gamma\sqrt{q}}{2\pi} \int_{-\infty}^{\infty} \chi(\|\mathbf{r}(q, p) - \mathbf{R}(\sigma, t)\|) \cos \sigma d\sigma. \quad (16)$$

## 8. Concluding Remarks

The structure developed in this paper provides a simple Hamiltonian description of the dynamics of an ideal fluid in a neighborhood of a vortex filament in terms of the self-induced motion of the filament. Thus the filament dynamics, which is induced by an infinite-dimensional Hamiltonian vector field, is the key ingredient in characterizing the fluid motion near the filament, and we plan to investigate this problem. Our illustrative examples reveal an apparent direct correlation between the symmetry of the initial filament and simplicity of the Hamiltonian structure. Further research is needed to fully comprehend this relationship.

To understand the global nature of the induced flow, one needs to study the evolving geometry and topology of the associated foliation in the large. This promises to be a formidable problem, but investigation of some interesting special cases would be instructive (cf. [1]). It is natural, and particularly relevant for applications, to ask if our Hamiltonian structure can be generalized to apply to two or more filaments. Some preliminary research of our own and related work such as in [2, 5, 10] suggest that the answer is affirmative for certain types of filament configurations.

## Acknowledgements

*The authors wish to thank Lu Ting for his perspicacious advice and assistance in all phases of the work on this paper. Conversations with Chjan Lim and Rupert Klein also proved helpful.*

## 9. References

- 1 ARNOLD, V.; KHESIN, B.: Topological Methods in Hydrodynamics. Springer-Verlag, 1998.
- 2 BLACKMORE, D.; KNIO, O.: KAM theory analysis of the dynamics of three coaxial vortex rings. *Physica D* **140** (2000), 321-348.
- 3 EBIN, D.; MARSDEN, J.: Groups of diffeomorphisms and the motion of an incompressible fluid. *Annals of Math.* **92** (1970), 102-163.
- 4 HASIMOTO, H.: A soliton on a vortex filament. *J. Fluid Mech.* **51** (1972), 477-485.
- 5 KLEIN, R.; MAJDA, A.; DAMODARAN, K.: Simplified equations for the interaction of nearly parallel vortex filaments. *J. Fluid Mech.* **288** (1995), 201-248.
- 6 LAMB, H.: Hydrodynamics. Dover, New York, 1932.
- 7 MARSDEN, J.; WEINSTEIN, A.: Coadjoint orbits, vortices and Clebsch variables for incompressible fluids. *Physica D* **7** (1983), 305-323.
- 8 MEZIĆ, I.; LEONARD, A.; WIGGINS, S.: Regular and chaotic particle motion near a helical vortex filament. *Physica D* **111** (1998), 179-201.
- 9 TING, L.: Studies in the motion and decay of vortices. *Aircraft Wake Turbulence*. Plenum, New York, 1971, 11-39.
- 10 TING, L.; KLEIN, R.: Viscous Vortical Flows. *Lect. Notes in Physics* **374**. Springer-Verlag, Berlin, 1991.
- 11 ZAKHAROV, V.: The algebra of integrals of motion of 2-dimensional hydrodynamics in Clebsch variables. *Func. Anal.*

*Addresses:* PROF. DENIS BLACKMORE, Department of Mathematical Sciences and Center for Applied Mathematics and Statistics, New Jersey Institute of Technology, Newark, NJ 07102- 1982 , USA.  
 PROF. OMAR KNIO, Department of Mechanical Engineering, The Johns Hopkins University, Baltimore, MD 21218, USA.

SAITO, SHIGERU

## VORTEX CAPTURING AND NOISE PREDICTION BY MOVING OVERLAPPED GRID METHOD

### 1. Introduction

The external noise of a helicopter is the most important issue in current helicopter technology. Specifically the reduction of the blade/vortex interaction noise is urgent subject in order for a civil helicopter to be spread widely. The capturing of the tip vortices in the noise analysis is the most important item because of its affection to sound pressure level. The moving overlapped grid method is developed in order to calculate rotor unsteady aerodynamics in hover and forward flights. This code is combined with the sound prediction codes using Ffowcs Williams and Hawkings formulation without quadruple term. The present combined noise analysis code is validated by the experimental data, which was conducted by Advanced Technology Institute of Commuter Helicopter (ATIC) at DNW in Netherlands. Comparison shows that the noise prediction was fairly good agreement with the data. Besides the calculated vortex parameters such as trajectory, core size, and strength of circulation are investigated by using present codes. The comparison of predicted tip trajectories shed from a rotor blade in hovering flight with well-known Kucurek's formula shows good agreement.

### 2. Experiment of tip vortex properties

Figure 1 shows the test set-up in the Low-speed Wind Tunnel facility at NAL[1]. Two different types of wing tip shape were used in order to investigate the effect on the tip vortex property. Three components of velocity vector were measured by 7-holes probe which is located just behind of the vortex generator. Using these velocity data, the core size and strength of circulation were estimated for two tip shapes. From this experiment, the core size and the strength of circulation for the tapered tip shape showed small values compared with that for the rectangular tip shape. In the noise prediction codes, mathematical models of tip vortex are often used to calculate the aerodynamic performance and trajectories of tip vortex for the blade/vortex interaction noise. In order to investigate the effectiveness of these models, several models were compared with test data. Figure 2 shows the comparison results for the case of rectangular tip shape. Among various vortex models, Scully vortex model is best fitted with present data. For the tapered tip shape case, Scully vortex model is still best fitted with the data. From these results, it is generally said that Scully vortex model is reasonable as tip vortex model.

### 3. Moving overlapped grid method(MOGM)

In the noise analysis, the free wake analysis based on the vortex theory is generally used in order to calculate the trajectory of the tip vortex. In this calculation, several empirical models, such as vortex core size and roll-up position of the tip vortex, are assumed. Therefore the accuracy of the calculation always depends on the decision of such parameters. To avoid these uncertain parameters, CFD technique is promising for the capturing the tip vortex. The moving overlapped grid method (MOGM) was developed to calculate the unsteady flow field around a rotor in the collaboration with Advanced Technology Institute of Commuter Helicopter Ltd. (ATIC)[2][3]. This CFD code was combined with the noise prediction code based on the Ffowcs Williams and Hawkings formulation without quadruple term. This CFD code has three grid systems (blade, inner and outer grid). The number of grid points is 17 millions. In order to get the convergence condition in the calculation, it took about one week by means of NWT at NAL.

### 4. Calculation results

Figure 3 shows the comparison of the tip velocity distribution between calculation and experiment. Experimental data was measured at 11 times of chord length behind the vortex generator. Calculation results shows that the about 1.5 times larger core size than that of the experiment. This is because the viscosity of the flow was not

taken into consideration in the calculation. Figure 4 shows the comparison of the prediction of noise waveform with experimental data. The experiment was conducted by the ATIC at DNW in Netherlands. The calculated results by MOGM show over-estimation in the magnitude compared with experimental data. This discrepancy between them were caused by the omission of the blade elastic deformation and the effect of the body itself on the rotor.

## 5. Conclusions

The wind tunnel test to measure the velocity field of tip vortex was conducted in the NAL Low-speed Wind Tunnel. These data were compared with calculation. The CFD code-named Moving Overlapped Grid Method was developed and combined with noise prediction code. From the investigation, the following conclusions are drawn.

1. The velocity field of a tip vortex is well modeled by the Scully vortex model.
2. CFD calculation results show the over-estimated prediction of core size of a tip vortex.
3. Noise predictions by the Moving Overlapped Grid Method show the over-estimation in the noise waveform. This is caused by the omission of the blade deformation and the effect of the body on the rotor.

In this calculation, Euler equation is used to predict the vortex property. Basically the viscosity of the flow is not taken into consideration. Therefore more sophisticate calculation should be necessary in the future.

## 6. References

- 1 YAMAGUCHI, A., SAITO, S., AOYAMA, T. AND OCHI, A.: Numerical Simulation around the tip region for several blades by Moving Overlapped Grid Method; Aerospace Numerical Simulation Symposium 99, June (1999)
- 2 OCHI A., SHIMA E., AOYAMA T. AND SAITO S.: Parallel Numerical Computation of helicopter Rotor by Moving Overlapped Grid Method; AHS International Meeting on Advanced Technology and Disaster Relief, Gifu, April 21-23 (1998)
- 3 OCHI A., AOYAMA T., SAITO S., SHIMA E. AND YAMAKAWA E.: BVI Noise Predictions by Moving Overlapped Grid Method; 55th American Helicopter Society Annual Forum, Montreal, May 25-27 (1999)

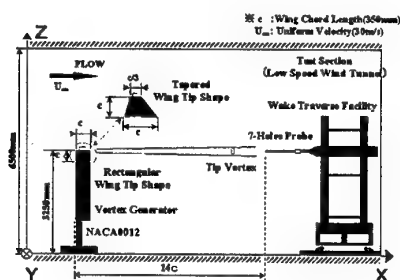


Figure1: Schematic view of test set-up

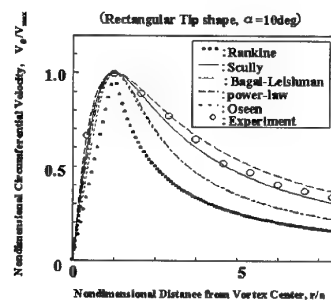


Figure2: Comparison of data with various models

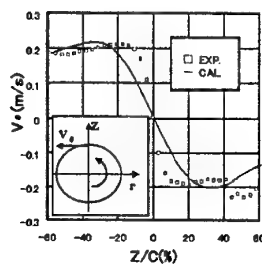


Figure 3: Comparison of tip velocity distribution between prediction and experiment

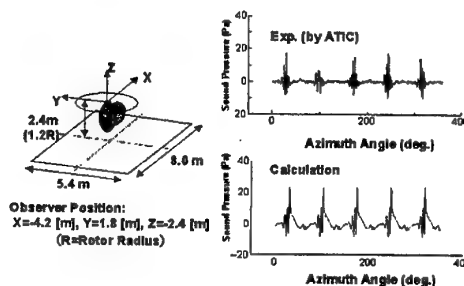


Figure 4: Comparison of the present calculation with the experimental data

Address: DR. SHIGERU SAITO, National Aerospace laboratory  
7-44-1, Jindaijihigashi-machi, Chofu, Tokyo, 182-8522, Japan

WAGNER, S.

## Calculation of Far Field Sound due to Blade-Vortex Interaction on a Helicopter Rotor

*The objective of the paper is to present an efficient tool for the prediction of far field noise due to blade - vortex interactions (BVI) on a helicopter rotor. These interactions occur mostly in low subsonic flight, e.g. a helicopter in descend. In this case, the method of retarded potentials can be used. When transonic effects at the tips of the advancing blade are present a coupled Euler-Kirchhoff method is applied. These two schemes allow the calculation of the acoustic signatures at arbitrary observer positions for rotors with usual blade motions. The theoretical results are compared with experimental data.*

### 1. Introduction

Noise generated by helicopters can give rise to problems concerning their public acceptance. Helicopters have many components that produce noise but it is usually dominated by the main rotor with the engine and the tail rotor being important additional sources. At this point only aerodynamic noise will be addressed. It is caused by several complex flow phenomena [12]. Inflow-turbulence noise is caused by the interaction of turbulence in the inflow with the leading edge of an airfoil. Trailing-edge noise and also tip noise is a by-product of any turbulent flow around blades. Typical noise spectra of helicopters (e. g. [9]) indicate that especially blade vortex interaction (BVI) noise plays a dominant role. High-speed impulsive (HSI) noise [1, 3] is another major noise source in high speed flight, when aerodynamic shocks are generated and radiate into space as sharp pressure peaks. These pressure gradients build up a wave front of spiral-like shape and are perceived as periodic noise. BVI noise will be described and how it can be simulated by theoretical/numerical methods. In case of low subsonic flight, e.g. a helicopter in descend, the method of retarded potentials is used. When transonic effects occur at the tips of the advancing blade a coupled Euler-Kirchhoff method is applied.

### 2. Calculation of Far Field Sound due to Blade-Vortex Interaction (BVI)

BVI noise is characteristic for helicopters and originates from the interaction of a rotor blade with the tip vortex of a preceding blade. This interaction leads to a sudden variation of the flow around the blade and therefore to a pressure fluctuation that is radiated away as sound. BVI is most important if the tip vortices pass very close to a blade and if the vortex axis is parallel to the blade. These conditions occur especially during descent flight.

The starting point for the prediction of BVI noise is usually the homogeneous linear wave equation [10]. In its most simple form it describes the propagation of sound waves through an acoustic medium at rest. The generation of sound is due to (1) the volume displacement of a solid body moving through the air, (2) the forces that this body exerts on the air, and (3) the non-linear momentum flux terms that may be caused by turbulence or by strong non-linearities in the flow field such as shocks. The first part is often referred to as thickness noise and is of monopole type. The second part is termed loading noise and is a dipole term, the third is mathematically a quadrupole term. It was Lighthill [7, 8] who recognized that this term is associated with the turbulent Reynolds stress tensor, i. e. the mixing of turbulent eddies which take place for instance in a jet [8]. This approach is called acoustic analogy.

In order to properly predict BVI noise the correct structure and position of the wake including the tip vortices must be known. There are several approaches possible [11]. In the present paper a linear theory, a non-linear theory and combination of both are used. The linear theory used is a vortex-lattice method for rotary wings (ROVLM). Zerle's computation [14, 15] included 240 time steps where the increment of the azimuth angle was gradually reduced from 15 degrees to 1 degree (Fig. 1). This solution allows the extraction of intensive BVI. In the present example they occur between 50 and 60 degrees azimuth angle in a low speed descendent flight where incompressible flow can be assumed. Zerle et al. [16] used the calculated pressure distribution in a so-called "retarded potential postprocessor" and calculated the sound received by an observer. The method [14, 16] is based on the linearized wave equation (no transonic effects, constant speed of sound  $a_0$ )

$$\nabla^2 \Phi - \frac{1}{a_0^2} \cdot \frac{\partial^2 \Phi}{\partial t^2} = 0$$

the extended GREEN's function

$$G = \frac{1}{r_{ret}} \cdot \delta \left( t - t_e + \frac{r_{ret}}{a_0} \right)$$

the potential function

$$\Phi_P = \underbrace{\frac{1}{4\pi} \iint_S -\sigma \cdot \frac{1}{r_{ret}} dS}_{\text{Source}} + \underbrace{\frac{1}{4\pi} \iint_S \mu \cdot \frac{\vec{e}_\mu \cdot \vec{r}_{ret}}{r_{ret}^3} dS}_{\text{Doublet}} + \underbrace{\frac{1}{4\pi} \iint_S \frac{\dot{\mu}}{a_0} \cdot \frac{\vec{e}_\mu \cdot \vec{r}_{ret}}{r_{ret}^2} dS}_{\text{Ratelet}} + \underbrace{\frac{1}{4\pi} \iint_W \mu_W \cdot \frac{\vec{n} \cdot \vec{r}_{ret}}{r_{ret}^3} dW}_{\text{Wake-Doublet}}$$

and the sound pressure calculation in the far field

$$p_S = \rho_0 \cdot \frac{\partial \Phi}{\partial t}$$

where  $r_{ret}$  is the retarded distance between the noise source and the observer,  $\sigma$  is the sink/source distribution (single layer potential) and  $\mu$  is the doublet distribution (double layer potential) of each panel, where the index  $W$  indicates the wake panels.  $\vec{e}_\mu$  marks the orientation of the local doublet. The blades have to be determined in a retarded position. The contribution of each blade to the sound received by an observer can thus be calculated (Fig. 1).

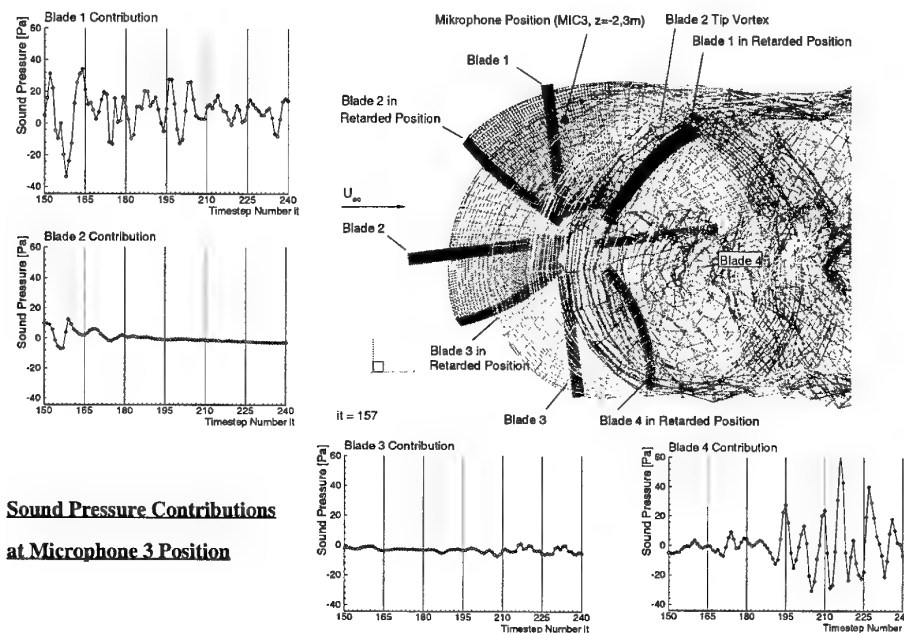


Figure 1: Sketch of vortex structure and sound pressure contribution of each blade at microphone 3, [16].

Since blade number 4 did not contribute that much noise in reality as calculated by theory because of the presence of a strong wake behind the rotor mast fairing which was not simulated this contribution was omitted and a good agreement with measurements was obtained (Fig. 2).

If transonic effects are present non-linear methods must be applied, e. g. the Euler equations. To resolve the structure of the tip vortices a very fine mesh and higher order algorithms in space have to be applied. Since, in addition, the distance between the noise source and the observer might be large a very big computational effort both in computer storage and speed is necessary. Therefore, Wehr [13] in a first attempt coupled ROVLM with the Euler code INROT of the institute (IAG) and calculated the unsteady pressure field induced by the rotor. The advantage of this procedure is that the wake of the rotor is predicted without any dissipation or diffusion because of the numerical solution of the vortex lattice method. The actual BVI is then calculated by the Euler procedure. However, the computational effort is still large. Thus, Algermissen [2, 3] coupled INROT with the Kirchhoff method to simulate the sound propagation to an observer.

Generally, when at transonic speeds so-called HSI noise is to be calculated the acoustic analogy strategy requires to evaluate the sound sources on the blade and in the flow field. The latter are necessary in order to model sound generation by aerodynamic shocks. Sound radiation from these sources is described by the Ffowcs Williams-Hawkings



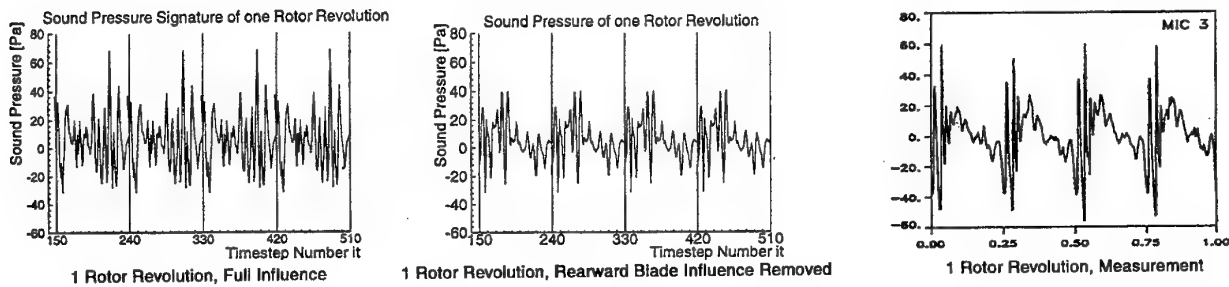


Figure 2: Comparison of calculated and measured sound pressures, [16].

(FW-H) equation [6]. Similar to the equation derived by Lighthill the FW-H equation is an exact combination of the continuity equation and the Navier-Stokes equations that results in an inhomogeneous wave equation. However, the FW-H formulation is much more suited for the treatment of noise from moving solid bodies as it displays the effect of volume displacement, forces, and non-linear terms as separate source terms on the right-hand side of the wave equation. The FW-H equation is the most general form of Lighthill's acoustic analogy [7] and is appropriate for the prediction of noise that is generated by helicopters where especially HSI and BVI noise occur.

The Kirchhoff formulation for moving surfaces is an alternative formulation that has widely been used in rotary wing aeroacoustics in the past [5]. However, there is a problem where to place the so-called Kirchhoff surface when it is moving [2, 3]. To compute the pressure at an arbitrary observer position both surface and volume integrations are necessary in the FW-H approach, whereas the Kirchhoff approach assumes validity of the homogeneous wave equation outside the Kirchhoff surface that is the boundary between the non-linear and linear calculation. For sound computation according to a Kirchhoff formulation integration over the two-dimensional Kirchhoff surface is sufficient. Since in transonic flow fields the local speed of sound is not constant and the velocity vector differs from free stream conditions sound propagates through such regions not in a perfectly uniform manner as assumed by the acoustic analogy. However, the Kirchhoff strategy makes the assumption of free flow sound convection only outside the Kirchhoff surface.

Both methods are usually applied in retarded time formulation, i. e. that different points of the spatial integration domain refer to different emission times but to the same observer time. For high-speed flight part of the sound sources move with supersonic speed. If the Kirchhoff surface is placed around these sound sources and rotates with the rotor blade parts of the Kirchhoff surface would also move with supersonic speed. In this case difficulties arise with acoustic analogy and Kirchhoff method. Multiple emission times can correlate to one observer time and singularities occur in the integration formulae. Farassat and Myers [5] presented a Kirchhoff formula for supersonically moving surfaces that can cope with these singularities. However, the procedure is more complex than the one for subsonic speeds. These difficulties were often circumvented in the past using a non-rotating Kirchhoff surface, enclosing the complete rotor tip plane. In forward flight the Kirchhoff surface was then assumed to move uniformly with the helicopter and a subsonic Kirchhoff formulation was used. Zibi et al. [17] showed that the position of the Kirchhoff surface can considerably influence the accuracy of the noise prediction. Algermissen and Wagner [1] found that for a hovering rotor the position of the Kirchhoff surface does not influence the results remarkably. They got good agreement with measurements. However, in forward flight the results did depend on the position of the Kirchhoff surface and the agreement with measurement was only fair. The numerical resolution of the code seemed to be too low (it was only third order accurate in computational space) to resolve the pressure fluctuations necessary to predict the emitted noise.

In a recent paper Brentner [4] re-examined the FW-H equation to demonstrate recent advances in rotor noise prediction and showed that the Kirchhoff formulation can lead to misleading results. In order to show the shortcomings of the Kirchhoff formulation he applied an embedding procedure similar to that used for the derivation of the FW-H equation to the wave equation by Farassat and Myers [5] and derived that way the Kirchhoff formulation for moving surfaces. Now he could compare the FW-H equation and Kirchhoff formulation and could that way show that for linear wave propagation several terms in the FW-H formulation would trend to zero and the FW-H equations and the Kirchhoff formulation would completely agree.

Finally in Figure 3 the prediction of acoustical pressure signals computed by Algermissen [3] are compared with measurements. The As mark positions of tip vortices whereas the INs indicate positions of inboard vortices. The agreement is only fair. A higher resolution of the Euler code seem to be necessary to resolve the pressure peaks that are induced by the tip vortices and that are responsible for BVI noise.

### 3. Conclusions

The paper presents two methods for the prediction of far field noise due to blade-vortex interaction (BVI). BVI noise usually occurs in low speed descent flight. A method derived from the linearized wave equation in combination with a post processor based on retarded potentials gives good results compared to measurements. If transonic effects occur this procedure is not applicable any more. An Euler/Kirchhoff method was then applied. In hover this method showed good results. In forward flight the problem occurred where to place the Kirchhoff surface since the results did depend on its position. Brentner showed with a new derivation of the Ffowcs Williams-Hawkings approach that this problem can be circumvented. In addition the resolution of the Euler method seemed to be too coarse. At least 4th order difference schemes in space seem to be necessary to resolve the influence of the vortex on the pressure peaks that are responsible for noise.

### 4. References

- 1 ALGERMISSEN, G. AND WAGNER, S.: Computation of High-Speed Helicopter Rotor Impulsive Noise by an Enhanced Kirchhoff Method. Procs. of the AHS Technical Specialists' Meeting for Rotorcraft Acoustics and Aerodynamics, Williamsburg, VA, 1997.
- 2 ALGERMISSEN, G. AND WAGNER, S.: Computation of Helicopter BVI Noise by Coupling Free-Wake, Euler and Kirchhoff Method. Procs AIAA/CEAS Aeroacoustics Conference, Toulouse, France, pp. 217-225, AIAA-98-2238, 2-3 June 1998.
- 3 ALGERMISSEN, G.: Eine Euler/Kirchhoff-Methode zur Berechnung von impulsivem Hubschrauberrotorlärm. Dissertation, Universität Stuttgart, 22.12.1998.
- 4 BRENTNER, K. S.: Modeling Aerodynamically Generated Sound: Recent Advances in Rotor Noise Prediction. AIAA Paper 2000-0345.
- 5 FARASSAT, F. AND MYERS, M. K.: Extension of Kirchhoff's Formula to Radiation from Moving Surfaces. Journal of Sound and Vibration, Vol. 123, No. 3, pp. 451-461, 1988.
- 6 FFWCS WILLIAMS, J. E. AND HAWKINGS, D. L.: Sound Generation by Turbulence and Surfaces in Arbitrary Motion. Philosophical Transactions of the Royal Society of London, Vol. 264, No. A 11e51, pp. 321-42, May 1969.
- 7 Lighthill, M. J.: On Sound Generated Aerodynamically; I. General Theory, Proceedings of the Royal Society of London Series A, Vol. 211, pp. 564-587, 1952.
- 8 Lighthill, M. J.: Jet Noise. AGARD Report 448, April 1963.
- 9 VAN DITSHUIZEN, J. C. A.: Helicopter Model Noise Testing at DNW- Status and Prospects. Procs. of the 13th European Rotorcraft Forum, Arles, France, 8-11 September, 1987.
- 10 WAGNER, S.; GUIDATI G. AND BAREISS R.: Wind Turbine Noise. ISBN 3-540-60592-4, Springer, Heidelberg, Mai 1996.
- 11 WAGNER, S.: On the Numerical Prediction of Rotor Wakes Using Linear and Non-Linear Methods. AIAA Paper 2000-0111.
- 12 WAGNER, S.: Aeroacoustics of Wind Turbines and Helicopter Rotors. Distinguished Keynote Address. Procs. of the 7th International Congress on Sound and Vibration, Garmisch-Partenkirchen, 4-7 July 2000.
- 13 WEHR, D.: Untersuchungen zum Wirbeltransport bei der Simulation der instationären Umströmung von Mehrblattrotoren mittels der Eulergleichungen. Dissertation, Universität Stuttgart, 10.03.1999.
- 14 ZERRLE, L.: Aerodynamische und aeroakustische Rotorberechnung unter Anwendung frei entwickelter Nachlaufwirbelschichten und retardierter Potentiale. Dissertation, Universität Stuttgart, 29.11.1999.
- 15 ZERRLE, L.; DOLD, R.; MICHL, TH. AND WAGNER, S.: Parallelization of a Vortex Lattice Rotor Code for Helicopter Aerodynamics. In: Parallel Computational Fluid Dynamics, Algorithms and Results Using Advanced Computers (P. Sciano, A. Ecer, J. Periaux, and N. Satofuka (Eds.)), Elsevier, Amsterdam, pp. 320-327, 1997.
- 16 ZERRLE, L. AND WAGNER, S.: Direct BVI - Sound Pressure Calculation by Using Aerodynamic Data and Application of Retarded Potentials. Procs. of the 25th European Rotorcraft Forum, Rome, Italy, Paper No. B9, 14-16 September 1999.
- 17 ZIBI, J.; ROUZAUD, O. AND POLACSEK, C.: Prediction of HIS Noise Using a Coupled Euler/Kirchhoff Method for a Helicopter in Hover Flight, Procs. of the 22nd European Rotorcraft Forum, Brighton, UK, Paper No. 49, 7-19 September 1996.

Addresses: SIEGFRIED WAGNER, Institute for Aerodynamics and Gas Dynamics, University of Stuttgart (IAG), Pfaffenwaldring 21, D-70550 Stuttgart, email: wagner@iag.uni-stuttgart.de

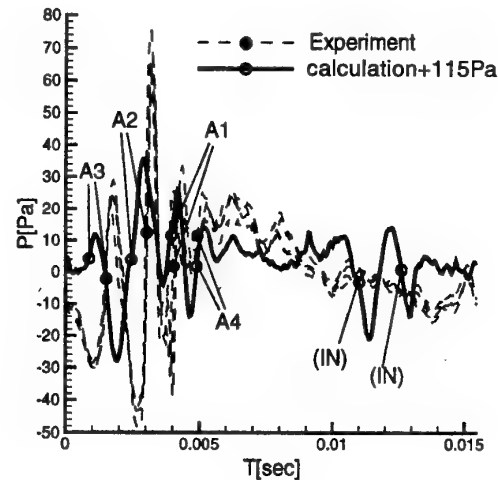


Figure 3: Measured and calculated acoustical pressure signals at an observer, [3].

BILICKI, Z.

## Wave Phenomena in Bubbly Liquid-Vapour Mixture (review)

*The mixture of a liquid and bubbles of its vapour is an excellent example of two-phase system where interesting wave phenomena take place. The occurrence of these phenomena in single-component two-phase mixtures can be attributed to the effect of strong thermodynamic non-equilibrium. During the flow of a two-phase mixture, mechanical non-equilibrium can also occur. Both types of non-equilibrium give rise to strong deformations of velocity and temperature fields in the liquid surrounding vapour bubbles. Specific features of the two-phase medium determine its wave properties, pronounced in its dispersive nature, meaning that the velocity of propagation of disturbances in the two-phase medium depends on their frequency. These above properties typical for the two-phase medium of low void fraction can be modelled using a concept of internal structure. In this concept, a model of continuum with the so-called operative parameters is assumed. The model is capable of predicting the phenomena such as: dispersive shock waves, pseudo-criticality, wave propagation and divergence between the hydrodynamic and thermodynamic pressure during flashing flow.*

### 1. Introduction

Real two-phase one-component system of a liquid and its vapour is distinctive by virtue of very complex physics, including thermodynamic and mechanical non-equilibrium. The thermodynamic non-equilibrium follows from the facts that the temperature of the liquid  $T_l$  is different from the temperature of the vapour  $T_g$ , i.e.  $T_l \neq T_g$ , and the chemical potential of the liquid  $\mu_g$  assumes different values than the chemical potential of the vapour  $\mu_l \neq \mu_g$ . The thermal non-equilibrium,  $T_l \neq T_g$  gives rise to energy transfer between the phases, whereas the chemical non-equilibrium,  $\mu_l \neq \mu_g$ , is a stimulus for the phase changes – evaporation and condensation, that can be considered the mass transfer between the phases. Mechanical non-equilibrium is pronounced in different local velocities of the liquid phase,  $w_l$ , and vapour phase,  $w_g$ ,  $w_l \neq w_g$ . This type of non-equilibrium is substantiated in the momentum transfer between the two phases. The exchange of mass, momentum and energy give rise to dissipation of energy as each of these processes is irreversible. The above mentioned processes take place at a finite rate, therefore it can be assumed that the return to the state of equilibrium is characterised by a finite relaxation time  $\vartheta$ . The two-phase liquid-vapour systems are peculiar for their relatively long relaxation times, especially during generation of the vapour phase, that is during flashing, where the relaxation time can be as high as 1 s, [1]. This compared to the relaxation times for gases or liquids, that are of the order of  $10^{-10}$ s, see Batchelor [2], brings realisation of the fact that the bubbly liquid-vapour system must be qualitatively different from the one-phase system. The differences are particularly pronounced in wave properties of bubbly flow.

### 2. Modelling two-phase flows

Real two-phase system is a discrete system. However, it is usually modelled as a continuum, [3]. In order to simplify the notation, let us limit ourselves to a one-dimensional model. The two-fluid model [4] will be assumed for consideration. Most models belonging to this class can be written in the form of the following set of equations

$$A_{ij}(\sigma_j) \frac{\partial \sigma_j}{\partial t} + B_{ij}(\sigma_j) \frac{\partial \sigma_j}{\partial z} = c_i(\sigma_j, z), \quad (1)$$

where  $\sigma$  is the state vector,  $t$  – time,  $z$  – coordinate along the channel. Imposing a small disturbance  $\delta\sigma$  on the state vector about undisturbed solution  $\sigma_l^0$  at a point  $l$ , so that  $\sigma = \sigma_l^0 + \delta\sigma$ , and making use of the analysis of small disturbances the following dispersive equation can be derived [5]

$$\det \left\{ A_{ij}(\sigma_l^0) - \frac{k}{\omega} B_{ij}(\sigma_l^0) + \frac{i}{\omega} \left[ \frac{\partial A_{im}}{\partial \sigma_j}(\sigma_l^0) \frac{\partial \sigma_m^0}{\partial t} + \frac{\partial B_{im}}{\partial \sigma_j}(\sigma_l^0) \frac{\partial \sigma_m^0}{\partial z} + \frac{\partial c_i}{\partial \sigma_j}(\sigma_l^0, z) \right] \right\} = 0. \quad (2)$$

Let us introduce the so-called phase velocity  $a_{ph} = \omega/k$ , where  $\omega$  is the disturbance frequency and  $k$  is the wave number linked to the wave length  $l$ ,  $k = 2\pi/l$ . The system defined by the set of Eqs. (1) can be described as wave-dispersive if the phase velocity depends on the disturbance frequency, that is  $a_{ph} = f(\omega)$ . This is possible if the

system is non-linear  $\frac{\partial A_{ij}}{\partial \sigma_i} \neq 0$ ,  $\frac{\partial B_{ij}}{\partial \sigma_i} \neq 0$  and when the source term  $c_i$  in Eqs. (1) depends on  $\omega$ , that is when  $\frac{\partial c_i}{\partial \sigma_i} \neq 0$ . The two-phase liquid-vapour bubbly flow is highly wave-dispersive, which is confirmed by numerous experiments, for example [6]. The dispersion is due to the thermodynamic non-equilibrium and the largest contribution follows from the term  $\frac{\partial c_i}{\partial \sigma_i} \neq 0$ . At the same time the two-phase flow is also characteristic for considerable damping of wave disturbances, a fact reflected in the model where the phase velocity  $a_{ph}$  is a complex number with a real and imaginary part.

### 3. Relaxation model

The most simple model that takes into account thermodynamic non-equilibrium is a relaxation model described by the following conservation equations of mass, momentum and energy [7].

$$\frac{\partial \rho}{\partial t} + \frac{1}{A} \frac{\partial \rho w A}{\partial z} = 0, \quad (3)$$

$$\rho \frac{\partial w}{\partial t} + \rho w \frac{\partial w}{\partial z} = -\frac{\partial P}{\partial z} - \frac{\tau C}{A} + \rho g, \quad (4)$$

$$\frac{\partial \rho u}{\partial t} + \frac{1}{A} \frac{\partial \rho u w A}{\partial z} = -\frac{P}{A} \frac{\partial w A}{\partial z} + \frac{\tau w C}{A}, \quad (5)$$

where  $w$  is the barycentric velocity,  $P$  – thermodynamic pressure,  $u$  – specific internal energy,  $\rho$  – density,  $A$  – cross-sectional area of the channel,  $C$  – channel circumference,  $\tau$  – wall shear stress,  $\tau = \frac{1}{2} f w^2 \rho$ , where  $f$  denotes the friction factor. The thermodynamic non-equilibrium is contained in the equation of state

$$\rho = \rho(u, P, x) \quad (6)$$

where  $x$  is the actual dryness fraction different from the equilibrium dryness fraction,  $\bar{x} = \frac{u - u'(P)}{u''(P) - u'(P)}$ , where  $u'$  and  $u''$  are specific internal energies corresponding to the saturation lines  $\bar{x} = 0$  and  $\bar{x} = 1$ . The above system of equations is augmented by a heuristic evolution equation, derived from the linear Taylor expansion of  $x$  about  $\bar{x}$

$$\frac{Dx}{Dt} = -\frac{x - \bar{x}}{\theta_x}, \quad (7)$$

where substantial derivative  $\frac{D}{Dt}$  is defined as

$$\frac{D}{Dt} = \frac{\partial}{\partial t} + w \frac{\partial}{\partial z}. \quad (9)$$

Quantity  $\theta$  is the relaxation time – a parameter that incorporates the internal structure of the two-phase system [8]. A more developed form of the evolution equation (7) was presented by Banaszekiewicz [9]

$$\frac{Dx}{Dt} = -\frac{x - \bar{x}}{\beta} - \frac{(x - \bar{x})^3}{\gamma} + \kappa \frac{\partial^2 (x - \bar{x})}{\partial z^2} \quad (9)$$

where  $\beta, \gamma, \kappa$  are constants that can be evaluated based on experimental investigations. The relaxation model (3-7) describes well flashing flow and explains a very interesting phenomenon of pseudocriticality [10], that is flow choking observed for Mach numbers as low as 0.2. However, the relaxation model fails to provide enough dissipation of energy especially for the description of propagation of large disturbances and stationary shock waves. In order to successfully model these phenomena, a model with the so-called operative viscosity  $\tilde{\eta}$  and operative heat conductivity  $\tilde{\lambda}$  was introduced.

### 4. Model with operative quantities

The operative viscosity  $\tilde{\eta}$  and operative heat conductivity  $\tilde{\lambda}$ , similar to turbulent viscosity in the turbulence models, substantiate strong deformations of the velocity and temperature fields in the carrier fluid at the vicinity of the dispersed phase. The model additionally takes into account the difference between the thermodynamic pressure  $P$  that is always positive, and the hydrodynamic pressure  $P^h$ ,  $P^\nu = P^h - P$ , where  $P^h = \frac{1}{3}(P_{xx} + P_{yy} + P_{zz})$ , which can assume either a positive or a negative value, and in general  $P^h \neq P$ . This model, described in detail in [11], has been found to describe well the propagation of large disturbances [12], structure of the stationary shock wave [13, 14], as well as the evolution of the pressure impulse during the rapid depressurisation of a liquid, [15]. Therefore, the model can be treated as the most general model capable of describing well the wave properties of the one-component two-phase bubbly system and can be written in the form as below

- mass conservation equation

$$\frac{D\rho}{Dt} + \rho \frac{\partial w}{\partial z} = -\frac{\rho w}{A} \frac{dA}{dz}, \quad (10)$$

- momentum conservation equation

$$\rho \frac{Dw}{Dt} = -\frac{\partial P^h}{\partial z} + \frac{\partial}{\partial z} \left[ \left( \frac{4}{3}\eta + \tilde{\eta} \right) \frac{\partial w}{\partial z} \right] - \tau \frac{C}{A} + \rho g \cos \alpha, \quad (11)$$

- energy conservation equation in the form of enthalpy,  $h$ , balance equation

$$\rho \frac{Dh}{Dt} - \frac{DP}{Dt} = \tau w \frac{C}{A} + (P - P^h) \frac{\partial w}{\partial z} + \left( \frac{4}{3}\eta + \tilde{\eta} \right) \left( \frac{\partial w}{\partial z} \right)^2 + \frac{\partial}{\partial z} \left( \tilde{\lambda} \frac{\partial T_s}{\partial z} \right) + q_w \frac{C}{A} \quad (12)$$

supplemented with the state equation

$$\frac{1}{\rho} = v = v(P, h, x) = v_l(P, x, h) + x[v''(P) - v_l] \quad (13)$$

together with the equations describing the evolution of the actual dryness fraction  $x$ , and pressure difference  $P^v = P^h - P$ . Based on the philosophy offered by the Thermodynamics of Irreversible Processes [15] there are 2 options for these equations. The first option comes from the Classical Irreversible Thermodynamics (CIT) and leads to the following couple of equations

$$\dot{x} = L_{11}(\mu_l - \mu_g) + L_{12} \frac{\partial w_i}{\partial z_i}, \quad (14)$$

$$P^v = L_{21}(\mu_l - \mu_g) + L_{22} \frac{\partial w_i}{\partial z_i}. \quad (15)$$

where  $\dot{x} = \frac{Dx}{Dt}$ . The second option is derived from the so-called Extended Irreversible Thermodynamics (EIT). EIT provides much richer evolution equations for the actual dryness fraction  $x$  and pressure difference  $P^v$  than CIT, namely

$$\Theta_x \frac{D\dot{x}}{Dt} + \dot{x} = L_{11}(\mu_l - \mu_g) + L_{12} \frac{\partial w_i}{\partial z_i}, \quad (16)$$

$$\Theta_P \frac{DP^v}{Dt} + P^v = L_{21}(\mu_l - \mu_g) + L_{22} \frac{\partial w_i}{\partial z_i}. \quad (17)$$

where  $L_{ij}$  are the fenomenological constants,  $\mu_l, \mu_g$  chemical potentials for liquid and vapour respectively. The following quantities appear in Eqs. (12), (13) are the dynamic molecular viscosity of the two-phase mixture  $\eta$ , the so-called operative viscosity  $\tilde{\eta}$ , molecular heat conductivity of the two-phase mixture  $\lambda$ , and its operative heat conductivity  $\tilde{\lambda}$ , wall shear stress  $\tau$ , heat flux to the channel from the outside  $q_w$ , channel cross-sectional area  $A$ , channel circumference  $C$ , acceleration of gravity  $g$ , angle between the flow direction and gravitational force  $\alpha$ , specific volume  $v$ , and  $v''$  - specific volume at the saturation line  $x = 1$ .

Eqs. (16) and (17) introduced into the conservation equations (10) - (12) constitute a model described by hyperbolic equations. This model predicts a finite value of the frozen velocity,  $a_f = \lim_{\omega \rightarrow \infty} a_{ph}$ . At the same time, the model contributes to the diffusion effects by second derivatives  $\frac{\partial^2 x}{\partial z_i^2}$ ,  $\frac{\partial^2 x}{\partial z_i \partial t}$ , and an extra term  $\frac{\partial^2 x}{\partial t^2}$ , which seems to be of importance for such a complex system as the two-phase mixture of a liquid and its vapour.

## 5. Conclusions

The two-phase liquid-vapour bubbly system is highly wave-dispersive and is characteristic for its long relaxation time, as a result of which first, wave disturbances are considerably damped, second, critical flow is connected with the equilibrium velocity, and then, there is a great impact of the bulk viscosity giving rise to a difference between the thermodynamic and hydrodynamic pressure. The properties of the considered two-phase system depend strongly on its internal structure. The effect of the structure can be expressed by the concept of operative viscosity and operative thermal conductivity.

### Acknowledgements

*The author would like to express his thanks to Doz. Dr. Martin Rein from DLR, the organiser of the minisymposium on "Compressible Multiphase Flow" for invitation to present this review talk.*

### 6. References

- 1 BILICKI, Z., KESTIN, J., PRATT, M. M.: A reinterpretation of the results of the Moby Dick experiments in terms of the non-equilibrium model; Transactions of the ASME, J. Fluid Engineering **112** (1990), 212–217.
- 2 BATCHELOR, G. K.: Introduction to Fluid Dynamics, Cambridge University Press, 1970.
- 3 BILICKI, Z.: Thermodynamic nonequilibrium in the two-phase system – a continuum with an internal structure; Archives of Thermodynamics **17** (1996), 109–134.
- 4 DELHAYE, J.M.: Basic equations for two-phase flow modeling; in Two-Phase Flow and Heat Transfer in the Power and Process Industries, eds A.E. Bergles, J.G. Collier, J.M. Delhaye, G.F. Hewitt, F. Mayinger, Hemisphere, 1981.
- 5 RAMSHAW, J.D., TRAPP, J.A. 1978: Characteristics, stability and short-wavelength phenomena in two-phase flow equation systems; Nucl. Sci. Engng. **66**, (1978), 93–102.
- 6 CHENG, L.-Y., DREW, D.A., LAHEY, JR. R.T. 1983: An analysis of wave dispersion, sonic velocity and critical flow in two-phase mixture; NUREG/CR-3372 Topical Report, Rensselaer Polytechnic Institute 1983.
- 7 BILICKI, Z. & KESTIN, J. 1990: Physical aspects of the relaxation model in two-phase flow; Proc. Royal Society London **A 428** (1990), 379–397.
- 8 BILICKI, Z., KWIDZINSKI, R. & MOHAMMADEIN, S.A.: Evaluation of the relaxation time of heat and mass exchange in the liquid-vapour bubble flow; Int. J. Heat Mass Transfer **39**, (1996), 753–759.
- 9 BANASZKIEWICZ, M.: Surface phenomena in nonequilibrium flashing flow; PhD thesis, Institute of Fluid Flow Machinery, Polish Academy of Sciences, Gdańsk, 1999.
- 10 DOWNAR-ZAPOLSKI, P., BILICKI, Z.: The pseudo-critical vapour-liquid flow; Archives of Thermodynamics **19** (1998) No. 3–4, 13–25.
- 11 BILICKI, Z.: Two-phase one-component system as a continuum with an internal structure; Proceedings of the International Symposium on Trends in Continuum Physics, TRECOP'98, eds B.T. Maruszewski, W. Muschik A. Radowicz, World Scientific, 1999.
- 12 BILICKI, Z., KARDAS, D., MICHAELIDES, E.E.: Relaxation models for wave phenomena in liquid-vapor bubble flow in channels; J. Fluids Engng., Trans. ASME **120** (1998), 369–377.
- 13 BILICKI, Z. & KWIDZINSKI, R.: Thermodynamic and mechanical non-equilibrium in flashing flow with shock wave in liquid-vapour mixture; Journal of Theoretical and Applied Mechanics **2** (1997), 35, 263–284.
- 14 KWIDZINSKI, R.: Experimental and theoretical investigation of the stationary shock wave in liquid-vapour mixture flow (in Polish); PhD thesis, IFFM PASc, Gdańsk, 1997.
- 15 BILICKI, Z.: Mathematical model of rapid depressurisation of a liquid; Proceedings of the Second International Conference on Heat Transfer and Transport Phenomena in Multiphase Systems, Kielce, May 18–22, 1999, 35–44.

*Addresses:* PROF. DR. HAB. ZBIGNIEW BILICKI, Institute of Fluid Flow Machinery, Polish Academy of Sciences, Gdańsk, Poland,  
e-mail: zb@karol.imp.gda.pl; <http://www.imp.gda.pl/struktura/o2/z4/ZB.htm>

LUTZ FRIEDEL

### Fluidodynamical critical two-phase (gas/liquid) flow state in non ideal flow geometries

During the emergency venting of pressurized, in most cases partially liquid loaded reactors across safety devices or in the course of the intentionally controlled depressurisation, resp., degassing of such production equipment across a throttle valve a two-phase flow can establish due to the occurring level swell. In the context of the equipment sizing for safety and economical reasons a maximum mass discharge, resp., critical velocity in the transfer piping system or device is an issue. Indeed, analytical methods for the critical two-phase mass flow prediction are only available for ideal nozzle or pipe flow. In principle, they look like simple derivatives of the ideal model used in (compressible) gas dynamics. Hence, rating tests for determination of the actual capacity of the flow controlling devices are necessary for an adequate sizing of the commercial devices. They are, however, time consuming and, therefore, prohibitively costly due to the large number of test variables as device size and type, fluid or mixture properties, initial subcooling, pressure, temperature, mass flow quality, valve lift, resp., opening, etc. are primary sizing parameters. Additionally, during the experiments the problem arise to define, resp., to fix the critical flow state. In contrast to the explanations in text books referring to the case of an ideal nozzle flow such a characteristic flow state does not establish in the device. This is, amongst others, the consequence of wall detachment of the flow, establishment of an inhomogeneous velocity field across the flow controlling cross section, local random cavitation or flashing/condensation incipience, multiple choking in the transfer piping system and corresponding back pressure change. As a consequence, the experimental maximum mass flow will depend, e.g., on the experimenter's pragmatism as well as capability and the metering quality installed in the test rig as well as, strictly speaking, it can only be valid for the individual specimen due to non avoidable manufacturing tolerances.

The engineering design practice is still characterized by particular solutions according to the respective limited experimental evidence as not at all a sizing convention is available. On the other hand, in all individual procedures, as a rule, the experimental or attributed maximum flow rate is expressed as, i.e., in single-phase compressible flow in terms of a corresponding one-dimensional ideal two-phase nozzle mass flow by introducing an empirical discharge coefficient of the valve. It is in this context invariably coupled to the respective flow model as well as specimen and its value should, according to the general physical understanding, remain below unity. Otherwise, the choice of the reference nozzle model is inappropriate. This is a common situation, as in two-phase flow an immense variety of critical mass flow models have been proposed. A common feature is, herewith, that for identical initial conditions different predictions are obtained though each model has been fitted to some experimental data and incorporates to a large extend theoretical aspects. Indeed, an analytically based pre-assessment and selection of the most promising two-phase nozzle flow model by using theoretical boundary values for comparison purposes is not feasible as these are not available.

The assigned discharge coefficient acts, in principle, as a global adjustment factor. In view of its definition it has no physical significance. Nevertheless, it is standing graphically for a variety of deficiencies of the reference flow model as dissipation, actual mixture state change, delayed phase change, non ideal fluid behaviour, incompleteness of the flow density definition, slip etc. In general, a lumped value of between 0.5 and 0.9 is obtained.

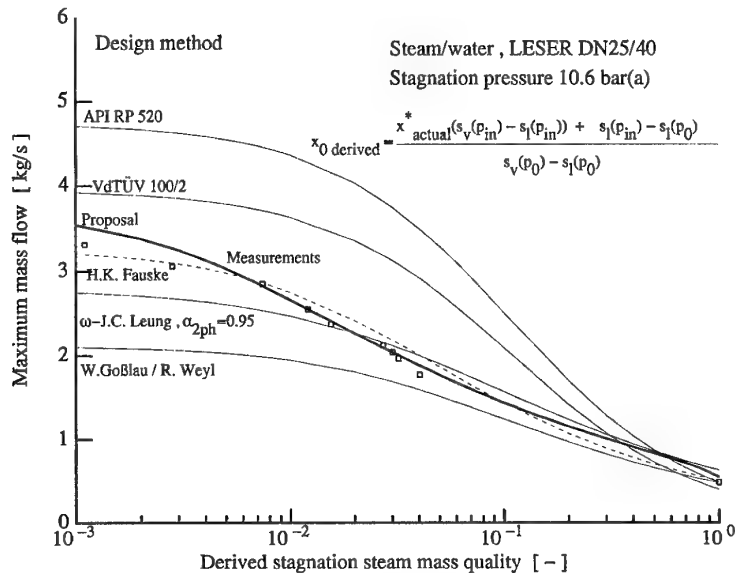
An example for such a pragmatistical approach is the recently proposed short-cut safety valve sizing method by T. Lenzing et al. [1]. The basis is formed by the isentropic homogeneous two-phase nozzle flow model, leading under all conditions to a discharge coefficient value in two-phase flow of less than unity. The essential features of the method are thermodynamic and fluiddynamic equilibrium between the phases. The other submodels included address the polytropic state change of the mixture, the boiling delay in subcooled or saturated flow, the degree of phase transfer and a liquid phase viscosity correction. As experimental pre-requisite for application at least one out of the two single-phase flow discharge coefficients must be available from the usual standardized (compulsory) device rating test. The weighed inclusion in the method is such that on approaching the limiting cases of the two-phase flow either the actual liquid or subcritical/critical gas/vapour flow capacity is predicted. The reproducibility of the underlying experimental results taken from the literature as well as produced in own tests is adequate, s. fig. 1 and 2. Indeed, the progress achieved in the reproductive accuracy against that of the formerly used methods is evident.

On concluding, it is fully recognised that the short-cut method also includes a combination of incomplete submodels and that the successful reproduction will be to some extent the consequence of mutual error compensation. Nevertheless, a moderate method extrapolatability with respect to fluids, pressure ranges and valve sizes is expected as physically reasonable intermediate key values are effective.

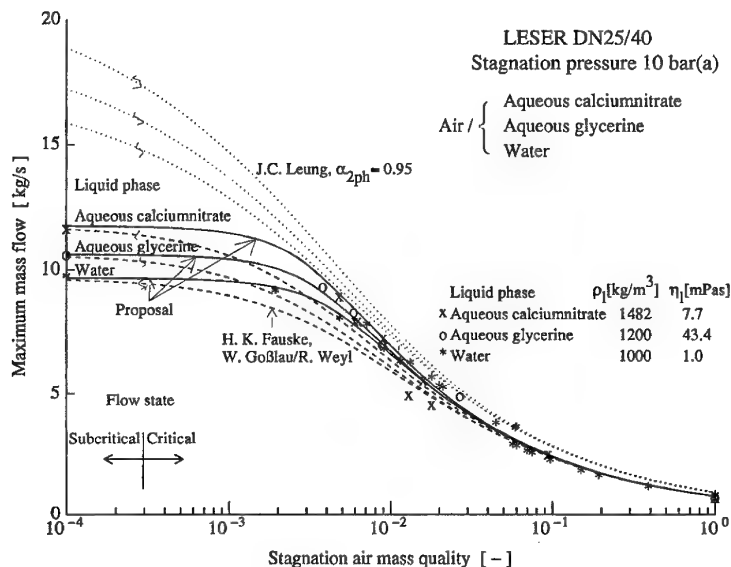
## 1. References

- 1 T. LENZING, L. FRIEDEL: Modelle für den über Vollhub Sicherheitsventile abführbaren Massenstrom bei Einphasen- und Zweiphasenströmung. Techn. Überw. 41 (2000) 5,6 und 7, 15/23.

Addresses: LUTZ FRIEDEL, Techn. Univ. Hamburg-Harburg, D 21071 Hamburg  
Germany



Experimental and predicted maximum dischargeable mass flow as a function of the steam mass quality for a LESER DN25/40 full lift safety valve



Experimental and predicted maximum mass flow as a function of the air mass quality and the liquid phase properties for a LESER DN 25/40 full lift safety valve



## Author Index

- Agarwal, R. K. 141  
Allgaier, R. 49  
Ams, A. 53  
  
Barkley, D. 81  
Bertram, A. 125  
Besson, J. 133  
Bilicki, Z. 155  
Bitterlich, S. 13  
Blackmore, D. 145  
Böhlke, T. 125  
Bontoux, P. 33  
Brocks, W. 133  
Bruns, J.-U. 105  
  
Chabanet, O. 133  
  
Denzler, J. 29  
  
Ertz, M. 57  
  
Friedel, L. 159  
  
Gaul, L. 45, 49  
Gersten, K. 73  
Grotjahn, M. 109  
Gumbsch, P. 137  
  
Heimann, B. 109  
Hellmich, Ch. 89  
  
Hintermüller, M. 17  
Hofstetter, G. 93  
  
Keiper, W. 49  
Kluwick, A. 77  
Knabner, P. 13  
Knio, O. 145  
Knothe, K. 57  
Kögel, R. 61  
Krause, R. H. 21  
Kunisch, K. 17  
  
Lackner, R. 89  
  
Macht, J. 89  
Mang, H. A. 89  
Margerit, D. 81  
Merkwirth, C. 113  
Müller, P. C. 117  
Müller, S. 61  
  
Oettli, G. 93  
  
Parlitz, U. 113  
Pawelski, H. 129  
Payret, R. 33  
Popp, K. 65, 105  
  
Raspo, I. 33  
Rudolph, M. 65  
Ruymgaart, F. 9  
  
Saito, S. 149  
Scheider, I. 133  
Schirmacher, R. 41  
Schmalfuß, C. 53  
Schmieg, H. 69  
Schuller, H. 97  
Schweiger, H. F. 97  
Serre, E. 33  
Solonnikov, V. A. 37  
Stark, R. F. 93  
Steglich, D. 133  
Sterpi, D. 101  
Stöbener, U. 45  
  
Ting, L. 85  
Tomlinson, G. R. 119  
  
Unbehauen, H. 121  
  
Van Dyke, M. 85  
Vielsack, P. 69  
  
Wagner, S. 151  
Wedig, W. 53  
Willner, K. 49  
Wohlmuth, B. I. 21  
  
Zumbusch, G. 25

# New book information

WILEY

## Multi-objective Optimization using Evolutionary Algorithms

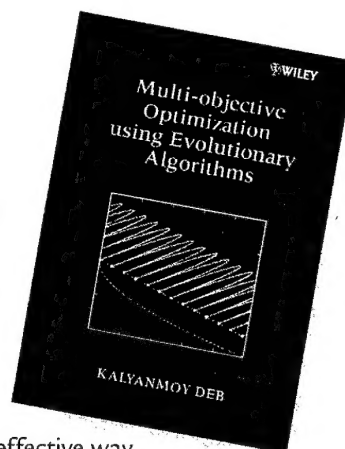
KALYANMOY DEB, Indian Institute of Technology, Kanpur, India

Evolutionary algorithms are relatively new, but very powerful techniques used to find solutions to many real-world search and optimization problems. Many of these problems have multiple objectives, which leads to the need to obtain a set of optimal solutions, known as effective solutions. It has been found that using evolutionary algorithms is a highly effective way of finding multiple effective solutions in a single simulation run.

- Comprehensive coverage of this growing area of research
- Carefully introduces each algorithm with examples and in-depth discussion
- Includes many applications to real-world problems, including engineering design and scheduling
- Includes discussion of advanced topics and future research
- Accessible to those with limited knowledge of classical multi-objective optimization and evolutionary algorithms

The integrated presentation of theory, algorithms and examples will benefit those working and researching in the areas of optimization, optimal design and evolutionary computing. This text provides an excellent introduction to the use of evolutionary algorithms in multi-objective optimization, allowing use as a graduate course text or for self-study.

0471 87339 X May 2001 496pp Hbk £60.00



For the latest information on forthcoming mathematics titles, please visit the Wiley Mathematics website at

<http://www.wiley.co.uk/mathematics>

For further information, or to place an order, please contact:

Zoë Mitchell, John Wiley and Sons, Baffins Lane, Chichester, West Sussex, PO19 1UD, UK  
Telephone: +44 (0) 1243 779777  
Fax: +44 (0) 1243 770154  
Email: [technology\\_uk@wiley.co.uk](mailto:technology_uk@wiley.co.uk)  
<http://www.wiley.co.uk>

### CONTENTS:

Prologue	Constrained Multi-Objective Evolutionary Algorithms
Multi-Objective Optimization	
Classical Methods	Salient Issues of Multi-Objective Evolutionary Algorithms
Evolutionary Algorithms	Applications of Multi-Objective Evolutionary Algorithms
Non-Elitist Multi-Objective Evolutionary Algorithms	Epilogue
Elitist Multi-Objective Evolutionary Algorithms	Bibliography

# Great News in 2002 - Price Reduction of ZAMM!

It is a great pleasure for Wiley-VCH to announce several fundamental changes in 2002 of

## **ZAMM** Zeitschrift für angewandte Mathematik und Mechanik Journal of Applied Mathematics and Mechanics

### 1. Price reduction of more than 650 DM!

Starting from 2002 the full price for institutional subscribers will be reduced from actual € 1628 to € 1300 p.a., which means you save € 328 (app. 650 DM) per year! Of course the number of articles in their usual high quality will not change.

### 2. Free access to the Electronic Publication of "Proceedings in Applied Mathematics and Mechanics"!

The proceedings of the GAMM-Conferences (starting with the Zurich 2001 GAMM-Conference) will be published in a new electronic journal "Proceedings in Applied Mathematics and Mechanics" which will be free of charge for GAMM members and subscribers of *ZAMM – Journal for Applied Mathematics and Mechanics*.

The *Proceedings in Applied Mathematics and Mechanics* will be part of Wiley InterScience ([www.interscience.wiley.com](http://www.interscience.wiley.com)) and available on CD-ROM.

### 3. Save time by Early View!

Articles of the electronic publication are available up to 75 days earlier than the printed version

For further information please contact us:

**WILEY-VCH Reader Service**  
P.O. Box 10 11 61  
D-69451 Weinheim, Germany

**Phone** +49 (0) 6201 / 606 147  
**Fax** +49 (0) 6201 / 606 117  
**E-mail** [subservice@wiley-vch.de](mailto:subservice@wiley-vch.de)

<http://www.wiley-vch.de/berlin>



Leichtweiß, K.

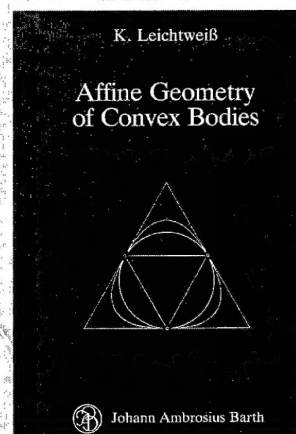
## Affine Geometry of Convex Bodies

1998. X, 310 pages with 10 figures.  
Hardcover.  
DM 148.-/€ 75.67/£Fr 131.-  
ISBN 3-527-40261-6

The theory of convex bodies is nowadays an important independent topic of geometry. The author discusses the equiaffine geometry and differential geometry of convex bodies and convex surfaces and especially stresses analogies to classical Euclidean differential geometry. These theories are illustrated by practical applications in areas such as shipbuilding. He offers an accessible introduction to the latest developments in the subject.

### From the contents:

1. Foundations of the equiaffine geometry
2. Geometrical meaning of equiaffine invariants for convex bodies
3. Inequalities
4. Affine characterization of special convex bodies
5. Equiaffine geometry of smooth convex bodies



PA0399

WILEY-VCH • P.O. Box 10 11 61 • 69451 Weinheim, Germany  
Fax: +49 (0) 62 01-60 61 84 • e-mail: [sales-books@wiley-vch.de](mailto:sales-books@wiley-vch.de) • <http://www.wiley-vch.de/>  
Prices are subject to change without notice.



## Contents

Preface	S 3
Table of Contents	S 5
Minisymposium 1	S 9
Minisymposium 2	S 13
Minisymposium 3	S 21
Minisymposium 4	S 29
Minisymposium 5	S 33
Minisymposium 8	S 41
Minisymposium 9	S 49
Minisymposium 10	S 73
Minisymposium 12	S 89
Minisymposium 13	S 105
Minisymposium 14	S 125
Minisymposium 15	S 133
Minisymposium 16	S 141
Minisymposium 17	S 155
Author Index	S 161